Analysis of cellular drivers of zebrafish heart regeneration by single-cell RNA sequencing and high-throughput lineage tracing

D i s s e r t a t i o n

zur Erlangung des akademischen Grades

Doctor rerum naturalium
(Dr. rer. nat.)

eingereicht an der
Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin

von

Bo Hu, M.Sc.

Präsidentin/Präsident
der Humboldt-Universität zu Berlin

Prof. Dr.-Ing. Dr. Sabine Kunst

Dekanin/Dekan der Lebenswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

Prof. Dr. Dr. Christian Ulrichs

Gutachter/innen

1. Dr. Jan Philipp Junker

2. Dr. Nikolaus Rajewsky

3. Dr. Pia Lundegaard

Tag der mündlichen Prüfung: 06.09.2021

Selbstständigkeitserklärung

Ich erkläre ausdrücklich, dass es sich bei der von mir eingereichten schriftlichen Arbeit mit dem Titel

Analysis of cellular drivers of zebrafish heart regeneration by single-cell RNA sequencing and high-throughputlineage tracing

um eine von mir selbstständig und ohne fremde Hilfe verfasste Arbeit handelt.

Ich erkläre ausdrücklich, dass ich sämtliche in der oben genannten Arbeit verwendeten fremden Quellen, auch aus dem Internet als solche kenntlich gemacht habe. Insbesondere bestätige ich, dass ich ausnahmslos sowohl bei wörtlich übernommenen Aussagen bzw. unverändert übernommenen Tabellen, Grafiken u. Ä. (Zitaten) als auch bei in eigenen Worten wiedergegebenen Aussagen bzw. von mir abgewandelten Tabellen, Grafiken u. Ä. anderer Autorinnen und Autoren (Paraphrasen) die Quelle angegeben habe.

Mir ist bewusst, dass Verstöße gegen die Grundsätze der Selbstständigkeit als Täuschung betrachtet und entsprechend der Prüfungsordnung und/oder der Allgemeinen Satzung für Studien- und Prüfungsangelegenheiten der HU (ASSP) geahndet werden.

Datum ……26.02.2021…… Unterschrift ……Bo Hu.....

Zusammenfassung

Das Herz eines Zebrafishs ist bemerkenswert, da es sich nach einer Verletzung vollständig regenerieren kann. Der Regenerationsprozess wird von Fibrose begleitet - der Bildung von überschüssigem Gewebe der extrazellulären Matrix (ECM). Anders als bei Säugetieren ist die Fibrose im Zebrafish nur transient. Viele Signalwege wurden identifiziert, die an der Herzregeneration beteiligt sind. Allerdings sind die Zelltypen, insbesondere Nicht-Kardiomyozyten, die für die Regulation des Regenerationsprozesses verantwortlich sind, weitgehend unbekannt. In dieser Arbeit haben wir systematisch alle Zelltypen des gesunden und des verletzten Zebrafischherzens mithilfe einer auf Mikrofluidik basierenden Hoch-Durchsatz-Einzelzell-RNA-Sequenzierung bestimmt. Wir fanden eine große Heterogenität von ECM-produzierenden Zellen, einschließlich einer Reihe neuer Fibroblasten, die nach einer Verletzung mit unterschiedlicher Dynamik auftreten. Wir konnten aktivierte Fibroblasten beschreiben und Fibroblasten-Subtypen mit einer pro-regenerativen Funktion identifizieren.

Darüber hinaus haben wir eine Methode entwickelt, um die Transkriptomanalyse und die Rekonstruktion von Zell-Verwandtschaften auf Einzelzellebene zu kombinieren. Unter Verwendung der CRISPR-Cas9-Technologie führten wir zufällige Mutationen in bekannte und ubiquitär transkribierte DNA-Loci während der Embryonalentwicklung von Zebrafischen ein. Diese Mutationen dienten als zellspezifische, permanente und vererbbare “Barcodes”, die zu einem späteren Zeitpunkt erfasst werden konnten. Mit maßgeschneiderten Analysealgorithmen konnten wir dann Stammbäume der sequenzierten Einzelzellen erstellen. Mit dieser neuen Methode haben wir gezeigt, dass im sich regenerierenden Zebrafischherz ECM-produzierende Zellpopulationen entweder mit dem Epi- oder mit dem Endokardium verwandt sind. Zusätzlich entdeckten wir, dass vom Endokardium abgeleitete Zelltypen vom Wnt-Signalweg abhängig sind.

Zusammenfassend beschreibt dieses Projekt einen allgemeinen Ansatz zur Analyse von durch Krankheiten oder Störungen ausgelösten Zelltypänderungen. Die Einzelzell-transkriptomanalyse liefert auch die Grundlage für die Identifizierung zusätzlicher pro-regenerativer Faktoren im Zebrafischherz.

Table of content

Summary

The zebrafish heart has the remarkable capacity to fully regenerate after injury. The regeneration process is accompanied by fibrosis - the formation of excess extracellular matrix (ECM) tissue, at the injury site. Unlike in mammals, the fibrosis of the zebrafish heart is only transient. While many pathways involved in heart regeneration have been identified, the cell types, especially non-myocytes, responsible for the regulation of the regenerative process have largely remained elusive. Here, we systematically determined all different cell types of both the healthy and cryo-injured zebrafish heart in its regeneration process using microfluidics based high-throughput single-cell RNA sequencing. We found a considerable heterogeneity of ECM producing cells, including a number of novel fibroblast cell types which appear with different dynamics after injury. We could describe activated fibroblasts that extensively switch on gene modules for ECM production and identify fibroblast sub-types with a pro-regenerative function.

Furthermore, we developed a method that is capable of combining transcriptome analysis with lineage tracing on the single-cell level. Using CRISPR-Cas9 technology, we introduced random mutations into known and ubiquitously transcribed DNA loci during the zebrafish embryonic development. These mutations served as cell-unique, permanent, and heritable barcodes that could be captured at a later stage simultaneously with the transcriptome by high-throughput single-cell RNA sequencing. With custom tailored analysis algorithms, we were then able to build a developmental lineage tree of the sequenced single cells. Using this new method, we revealed that in the regenerating zebrafish heart, ECM contributing cell populations derive either from the epi- or the endocardium. Additionally, we discovered in a functional experiment that endocardial derived cell types are Wnt signaling dependent.

In summary, by combining single-cell lineage analysis and single-cell transcriptomics, this project provides a general approach for dissecting cell type changes during diseases or tissue perturbations. The single-cell transcriptomics analysis also provides the basis for identifying additional pro-regenerative factors in the zebrafish heart.

Introduction

Zebrafish heart: a regeneration-competent organ

The zebrafish heart is a muscular organ that consists of two chambers: the atrium and the ventricle. The chamber walls are divided into three cellular layers: the endocardium as the endothelial cell layer on the inner side of the chamber with contact to the blood flow, the epicardium on the outer side of the heart, and the myocardium (the heart muscle) between these two layers (Fig. 1a). Like tissues of every multicellular organism, the zebrafish heart consists of many different cell types which have well-defined morphology, function and developmental origin. Some examples are cardiomyocytes (CMs), endocardial cells, epicardial cells, fibroblasts, and immune cells.

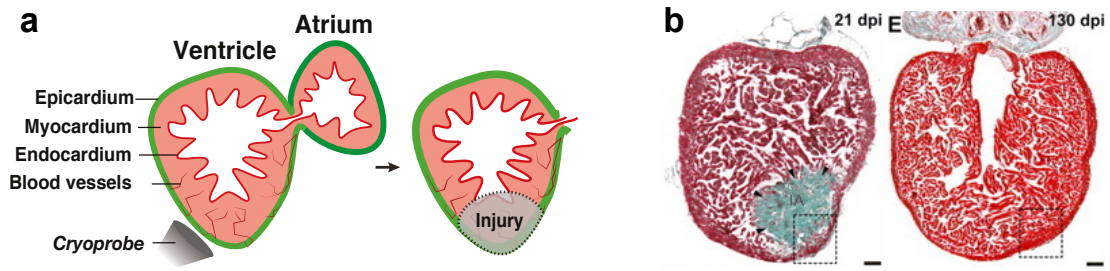


Figure 1. Zebrafish heart regenerates after cryo-injury.

a. A depiction of the layered structure of the zebrafish heart and the cryo-injury procedure. The cryoprobe causes a local injury. b. Histological staining of a fibrotic heart after cryo-injury and a fully regenerated heart. After cryo-injury, the fibrosis is visible as blue stain (Trichrome staining). It is cleared and replaced by heart muscle tissue after regeneration. Pictures modified from González-Rosa et al. (2011), Development.

The zebrafish heart, like other zebrafish organs, can fully regenerate after injury. Different approaches of injury and perturbation of the cardiac tissue have been developed for the zebrafish, such as ventricular apical resection and genetically targeted cell ablation[1,2]. Cryo-injury[3], especially, was described to mimic aspects of myocardial infarction in mammals (Fig. 1a). This property of the zebrafish heart and the available experimental approaches make the zebrafish an excellent model system for heart regeneration with translational potential. The regeneration process after cryo-injury is accompanied by fibrosis - the formation of excess extracellular matrix (ECM) tissue, at the injury site. Unlike in mammals, the fibrosis of the zebrafish heart is only transient[3] (Fig. 1b). In previous studies, lineage tracing experiments using transgenic markers have shown that the damaged heart muscles regenerate via dedifferentiation

and proliferation of cardiomyocytes[4,5]. After injury, they undergo limited dedifferentiation, disassemble their sarcomeric structure, and detach from one another[5].

It has been shown in previous studies that a complex interplay of molecular signals from different sources regulates the regeneration process. For example, important signals were found in the epi- and endocardium[24,25,26,27]. Immune cells, especially macrophages, were shown to be involved in the first stages of injury response and regeneration[28,29]. Cardiac fibroblasts, a major cell population in the heart, are activated and proliferate upon injury. They are the main contributors to the fibrosis and their ablation leads to reduced cardiomyocyte proliferation[29]. All together, these cell types create an environment that is beneficial for the regeneration of cardiomyocytes.

However, it remains unknown which cell types exactly are recruited to or generated at the site of injury in the regenerating heart. Due to the plethora of different cell types, it is also difficult to dissect the potentially complex pathways in play[30]. Our understanding of the regeneration process of the zebrafish heart would hence highly benefit from a thorough analysis on the cellular level.

Single cell RNA-seq: a powerful technology

Profiling of single cells on the transcriptome-wide level has recently become possible due to the emergence of single-cell RNA sequencing (scRNA-seq)[6-10]. The starting tissue material is first dissociated into a suspension of single cells (Fig. 2a), and their transcriptome is then captured one by one. Several methods for high-throughput scRNA-seq have been published[11-15], and can be exploited for time and cost efficient processing of large numbers of single cells. These methods either rely on automated pipetting robotics[11], combinatorial design of cell indexing[12], or droplet-based microfluidics techniques[13-15] (Fig. 2b), and the libraries are analyzed by high-throughput next generation sequencing. The sequenced singe cells are then clustered by their transcriptome similarity, for example using graph based clustering algorithms[16,17]. This results in a clustering of cells with similar transcriptome profiles (depicted in Fig. 2c). The genes that are differentially expressed in each cluster can then be calculated, resulting in an unbiased list of genes that define a certain cluster.

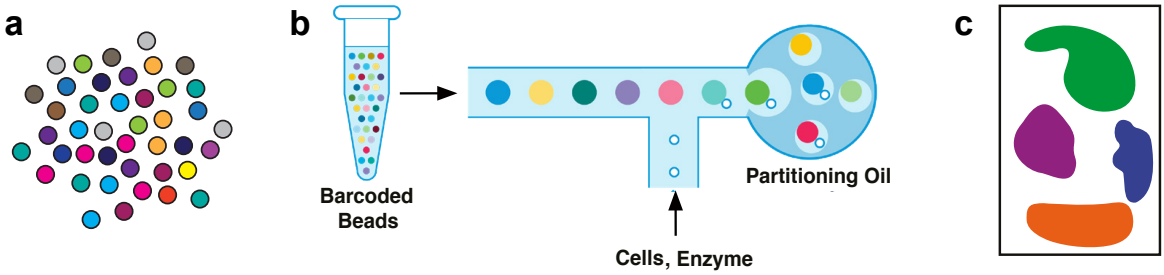


Figure 2. Workflow summary of microfluidics based scRNA-seq.

a. Tissue samples are dissociated into single cells, here depicted by colored dots. b. Single cells are captured in droplets with barcoded beads using a microfluidics device. c. After sample processing, sequencing and data processing, single cells with similar transcriptome can be visualized as clusters, here shown as colored patches.

Experiments following these approaches have allowed for systematic and unbiased identification of different cell types and their marker genes across different organisms, also in the developing and adult heart[18,19,20]. For example, there has been a recent study presenting a large-scale single-cell atlas of the healthy human heart[21], as well as a study on cell type changes in the mouse after myocardial infarction[22]. A single-cell analysis of sorted zebrafish cardiomyocytes also described their metabolic changes during the regeneration process[23]. However, no systematic analysis of the zebrafish heart has yet been published.

Single-cell lineage tracing cannot be combined with scRNA-seq so far

To fully understand how cell types change in number, and to reveal their process of activation after injury, it is also crucial to understand their origin. Knowing their lineage would allow us to determine the mechanisms that generate new cell types, and is hence an important task in studying regeneration.

An example here are the cardiac fibroblasts. Despite considerable experimental efforts, the lineage origin of cardiac fibroblasts is still controversial. Literature suggests several sources for them: activation of resident fibroblasts of epicardial origin, de novo generation of endocardial derived fibroblasts, or fibroblasts derived from hematopoietic cells[31]. Since not all of them respond equally upon injury, it has also been suggested that there could be different types of fibroblasts in the zebrafish heart[31]. Combining the transcriptome and lineage information on the single-cell level would enable us to find out where a transient or an activated cell type originates from.

However, broadly applicable experimental methods that analyze lineage and transcriptome information in single cells at the same time are lacking. So far, approaches for cell lineage analysis with single-cell resolution are either generally not capable of additionally detecting the transcriptome, as they are visual techniques that rely on microscopy, like the brainbow[32] approach and the single-cell live tracking technique[33,34], or they are limited to a certain types of tissue, like the retroviral labeling of cells[35-37], that requires ex vivo infection and cell transplantation and is only easily applicable to the hematopoietic system.

Project Objective

The aim of the PhD project was hence twofold.

First, to establish a new single-cell lineage tracing method. We wanted to use the capacity of the genome for information storage in order to bring lineage tracing to the high-throughput level, with a new single-cell approach that is time and cost efficient and can be combined with single-cell transcriptome profiling. This approach will be broadly applicable beyond our own experimental settings, for example in developmental biology, stem cell research and cancer biology.

Second, to utilize the enormous multiplexing capacity of high-throughput RNA sequencing in combination with the new lineage tracing approach to study all cell types, potentially also transient ones, in the context of zebrafish heart regeneration. We hypothesized that especially fibroblasts, located at the injury area, could have crucial functions for regeneration. We intended to characterize individual cell types after an injury and compare them with those we find in the healthy zebrafish heart. This would allow us to identify changes in the cellular composition of the zebrafish heart in the context of regeneration, and potentially find transient cell types. By describing the dynamics, location, origin, and signaling factors of these cell types, we aimed to identify potential pro-regenerative cell types and understand their mode of activation.

Finally, we wanted to contribute to a better understanding of the regeneration process of the heart of the zebrafish, the preeminent model for vertebrate heart regeneration.

Experimental strategy

Design of the lineage tracing strategy

Some key features should apply to a lineage tracer in general: First, it should not change the behavior of the marked cell, its progeny, and its surrounding. Second, the tracer must be passed on to all progeny of the cell and should be retained over time. Third, it should be never transferred to unrelated cells. In addition, our desired lineage tracer needs to be obtainable by sequencing, so it can be combined with scRNA-seq technologies.

Our strategy was to utilize the CRISPR-Cas9 system to induce heritable, non-editable mutations in the genome of the zebrafish (Fig. 3a). Once Cas9 induces double-strand

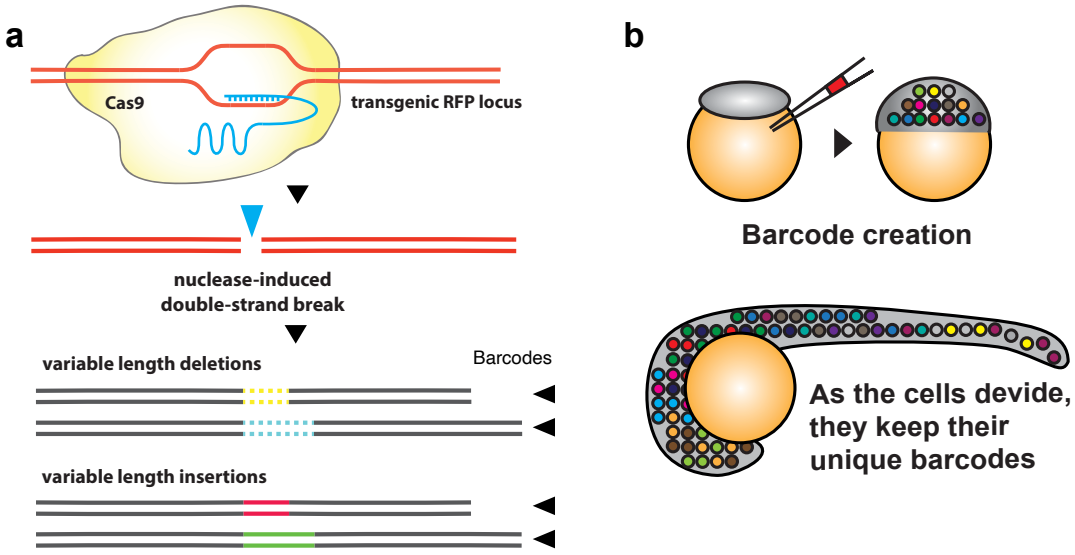


Figure 3. Overview of the barcode creation process.

a. Cas9 (yellow) and sgRNA (blue) create insertions or deletions in an RFP transgene. These genetic scars can be used as lineage barcodes. **b.** The Cas9 and sgRNA are injected into the zygote at one cell stage. As the cells divide, barcodes will be created and passed on to progenies (depicted by different colors). The RFP transgene and the barcodes are ubiquitously expressed and can be read out together with the transcriptome by scRNA-seq at a later stage.

breaks at its target site, a short deletion or insertion (indel), varying in length and position, is generated after the repair by non-homologous end joining. We injected Cas9 and sgRNA into zebrafish embryos at the 1-cell stage. As the cells divide, these

indels will be passed on to their progenies and new indels will be created within those, until either the Cas9 activity diminishes or its targets are exhausted. These indels subsequently label single cells and its progenies in the zebrafish (Fig. 3b) like barcodes, which fulfill the previously mentioned requirements for a desired lineage tracer:

1. It does not change the behavior of the marked cell: The loci we targeted for the induction of mutation were RFP-transgene integrations in the fish line Zebrabow-M[32]. It is very unlikely that the mutation in a transgene will affect the regular development of the zebrafish.

2. It is permanent: Once a barcode is created in a cell, the same target site cannot be cut again by Cas9, making the barcode permanent in that cell, as well as in all progenies of that cell.

3. It is not transferred to unrelated cells: For the readout, we would only select barcodes that have a low probability to be created and only would sequence one individual fish at a time, making it very unlikely to find the same barcode in unrelated cells.

4. Most importantly, The *ubi*-promotor driven RFP-transgene is ubiquitously expressed and therefore can be captured with the transcriptome in all cells by subsequent RNA sequencing methods.

It has previously been reported that this line has 16-32 identical integrations of RFP (in the F0 generation). Hence, with the same guide-RNA, we were able to create several different mutations in the same cell. With these mutations acting like lineage barcodes, they allow for the reconstruction of lineage trees using overlapping lineage barcodes from single cells.

Single-cell RNA sequencing platform

We aimed to use high-throughput scRNA-seq methods to simultaneously detect the transcriptome and barcodes in single cells. The goal was to profile thousands of cells in the regenerative process from the whole zebrafish heart. For this purpose, we compared four scRNA-seq platforms: plate-based single cell sorting (modified CEL-seq protocol[9]), Drop-seq[13], inDrops[14] (commercialized by the company 1-cell-bio), and Chromium[15] (commercialized by the company 10X Genomics, PN-120233). The latter three platforms all utilize a droplet-microfluidics device to achieve the mixing of single cells with barcoded primer beads in a single droplet. The single-cell sorting method was low in throughput compared to other methods. With the inDrops system, it was difficult to achieve reproducible replicates, since it only runs one sample at a time, and requires an extended manual reloading process for the reagents after each sample. The recovery rate of cells was relatively low in the Drop-seq system, which was detrimental for our lineage tracing approach, since our input material of cells was restricted to one zebrafish heart to avoid overlapping of barcodes in different fish. We came to the conclusion that the Chromium platform was the most suitable for our approach. It allowed for more replicates (up to 8 at a time) and less batch effects, and the cell recovery rate per sample was over 50%.

Targeted amplification of lineage barcode sequences

Since there is a strong bias towards the 3' end of the mRNA in the scRNA-seq method we were using, the RFP specific sgRNA for lineage barcode induction was designed to target the 3´ end of the RFP gene sequence. To ensure the effective capturing of the barcode sequences, we chose to amplify the sequences in a targeted manner using specific primers. The primers for barcode sequence amplification were designed in such a way that the Cas9 cutting site is positioned in the middle of the sequencing read. This made sure that the read will cover the whole barcode sequence, unless the barcode created was longer than the sequencing read itself (150 base-pairs).

We used the cDNA, which is generated during the protocol of the Chromium workflow[15], as the starting material for lineage barcode amplification. To be able to keep the sequence features that Chromium uses to distinguish single cells (cell barcodes) (Fig. 4) and transcripts (unique molecular identifiers[38]), one of the amplification primers needs to bind to the adapter sequence of the cDNA. This left us only one target specific primer for RFP sequence. We used a nested PCR approach (2 rounds of PCR with two different RFP specific primers) to circumvent the potential off-targets in the amplification process. Later, we were able to match the lineage barcodes to the transcriptome of every sequenced cell.

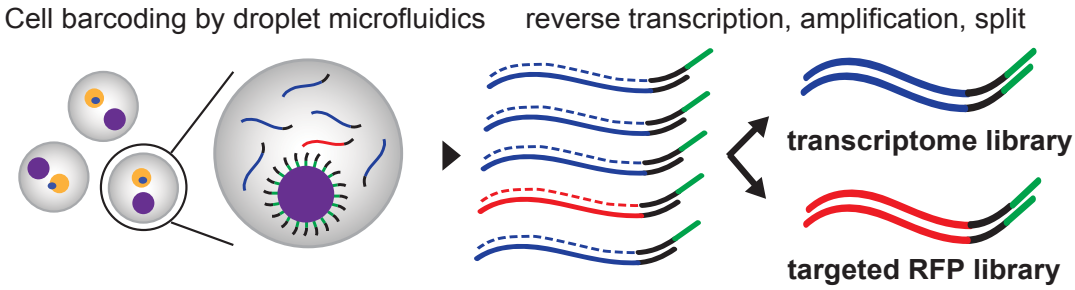


Figure 4. Generation of targeted barcode library.

The transcriptome of single cells is captured in single droplets with barcoded beads (barcode sequences depicted in green). The cDNA of single cells (shown in blue) is barcoded during reverse transcription. Afterwards, the lineage barcode sequences are amplified with targeted primers from the RFP transgene sequences (here shown in red) within the cDNA.

Experimental workflow

The experimental timeline is depicted below. The fertilized oocytes were injected with Cas9 and sgRNA at the one cell stage. Lineage barcodes were hence created very early in the development of these embryos. We raised the fish to adulthood (~6 months) and then either analyzed the healthy organs (not shown here), or performed heart cryo-injury experiments. To analyze the process of regeneration, we collected injured hearts after 3, 7 or 30 days post injury (dpi). These are time points where

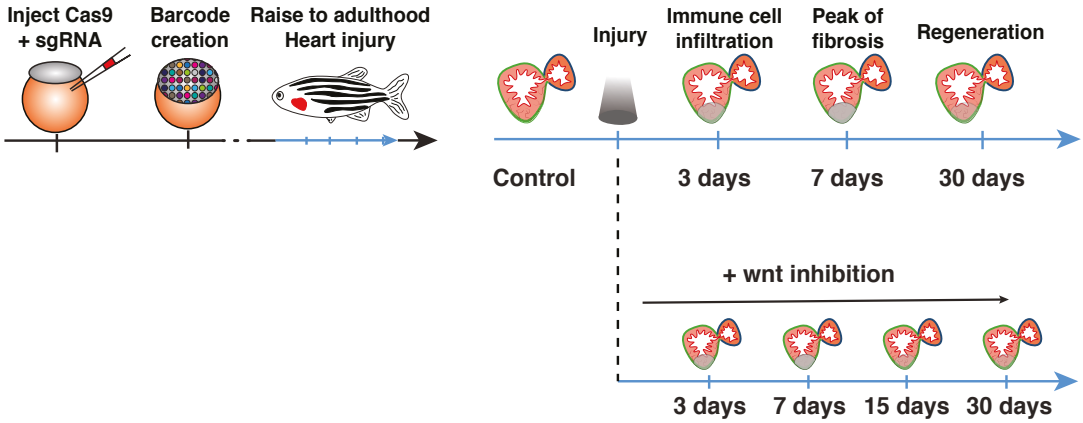


Figure 5. Experimental workflow.

Oocytes are injected at the one cell stage with Cas9 protein and sgRNA. Barcodes are created during the embryonic development. The heart injury experiments are performed in adult fish (timeline shown as blue arrows). The zebrafish hearts are analyzed either as a healthy control or at the time points 3, 7 and 30 dpi. After Wnt-inhibition, the hearts were studied additionally at 15 dpi.

major events take place during regeneration. At 3 dpi, the immune cell infiltration is at its maximum, and the peak of fibrosis was found at 7 dpi. After 30 days, the regeneration process should be almost completed. As we sequenced the zebrafish hearts, we would expect to see transcriptome changes of different cell types that reflect the different stages of regeneration. Since we would have the lineage information of those cell types, we would also be able to observe events like transdifferentiation and potentially the recruitment of cells from outside the heart. To investigate the effect of a broad perturbation on the cell types of the zebrafish heart during the regenerative process, we also designed a Wnt inhibition experiment using the canonical Wnt-antagonist IWR-1. We kept the inhibition for 30 days after the injury and collected hearts at 3, 7, 15 and 30 days post injury (Fig. 5).

Results

First, we developed and characterized the new sequencing-based lineage tracing method. The results of these experiments and their analysis described here were published in the paper Spanjaard, Hu et al. (2018), Nature Biotechnology[39].

Cas9/sgRNA injection was sufficient to create indels

As a first-line control for the effectivity of Cas9/sgRNA injection and the experimental variability, we used the fact that the mutations we introduced with our approach will lead to a loss of RFP fluorescence. 24 hours after the injection at 1 cell stage, we could observe the loss of red fluorescence in a significant fraction of the embryos compared to uninjected controls (Fig. 6a). The fish remaining brightly fluorescent were discarded from further experiments. After sequencing, we found up to 75% reduction of wild-type sequences in injected embryos. To rule out unspecific effects of the Cas9 injection, we injected Cas9 in combination with a random sgRNA. There, only wild-type RFP sequences were detected (Fig. 6b).

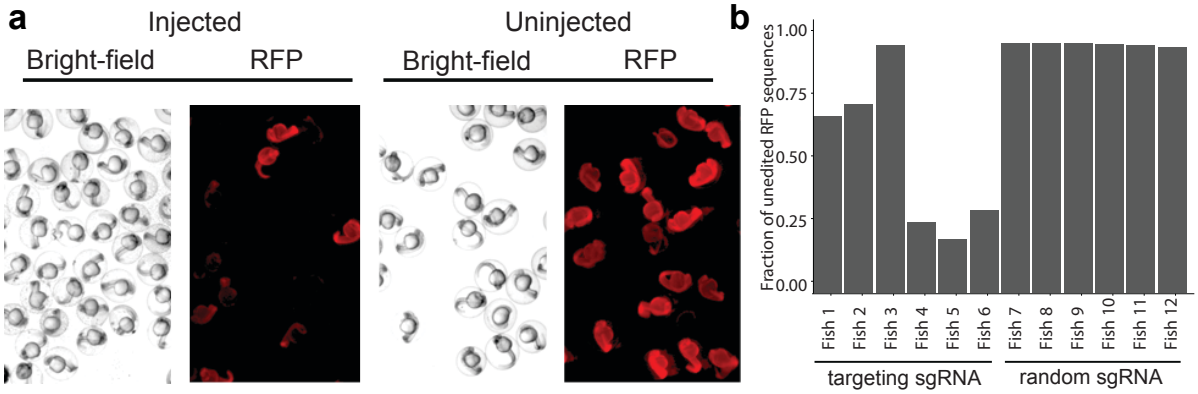


Figure 6. Cas9 and sgRNA injection creates indels in the RFP transgene locus.

a. Upon injection, RFP fluorescence decreases in the majority of the embryos, indicating mutations being created in the RFP sequence. Bright embryos after injection are excluded for further analysis. **b.** Sequencing mutations (indels) are detected by bulk sequencing in Cas9-targeted RFP locus.

High diversity of indels was created within 10 hours

To access the diversity and the dynamics of indel creation, we sequenced the DNA and mRNA of 32 embryos after Cas9/sgRNA injection at 1, 2, 3, 4, 6, 8, 10 and 24 hours post fertilization. We showed that the indels can be detected with similar efficiencies from both mRNA and DNA (Fig. 7a). This confirmed the possibility to detect the DNA mutations on the mRNA level, which will allow for the combination of transcriptome and barcode readout later.

We observed that mutation events happen within the first 10 hours of embryonal development (Fig. 7b). The mutations created in this way were highly diverse, including various deletions and insertions (Fig. 7c). We pooled all the mutations detected in all embryos and determined the probability for each of them to occur. We found out that some of them were created very often, while some of them were created rarely, with probabilities ranging between 10^{-6} and 10^{-1}(Fig. 7d).

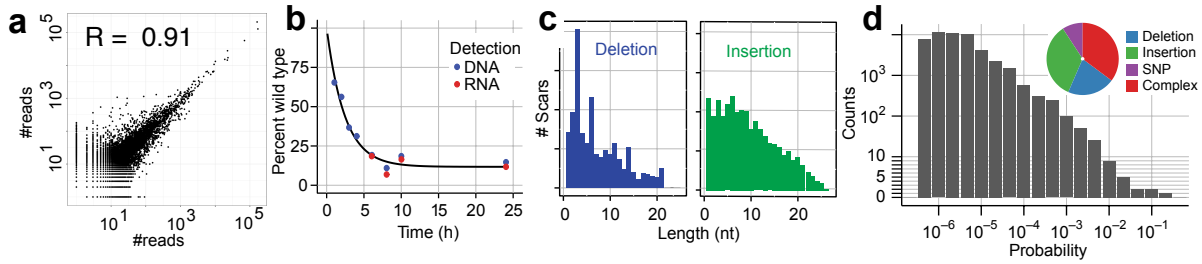


Figure 7. Characterization of mutations induced by Cas9 injection.

a. Correlation of scar abundances detected in bulk on the DNA and RNA level. Each dot represent a certain sequence at the mutation site. **b.** Dynamics of mutations created, as measured on the DNA and RNA level in bulk samples of pooled embryos, with exponential fit. **c.** Length distributions for detected deletions and insertions. **d.** Probability distribution of mutations, measured in bulk experiments on the DNA level. Pie chart shows fractions of different types of scars (SNP: single nucleotide polymorphism).

Detection of transcriptome and lineage barcodes in single cells

In the next step, we proceeded to use high-throughput scRNA-seq to detect the transcriptome and lineage barcodes in single cells at the same time. We first chose to study dissociated cells of the whole larvae at 5 days post fertilization (dpf). Since the cell number of a developing larva is relatively low, we were able to capture cells from all body parts of the zebrafish in one sample. We used the Chromium platform and its pipelines (10X genomics, see experimental strategy) for single-cell capturing, data demultiplexing, and transcriptome mapping. We used the R package seurat[16,17] for filtering, processing, and visualization (see Experimental Strategy and Methods).

After clustering of the single-cell transcriptome data, we could identify 70 clusters of cells. Based on their differential gene expression, we could identify cell types from all three developmental layers: the ectoderm, the mesoderm and the endoderm. (Fig. 8 and Fig 9 for overview of the clustering, for detailed t-SNE and gene lists see also Additional Figure 1 and Additional table 2). From seven biological replicates, we sequenced more than 70 000 cells with ~700 genes detected per cell on average. Together with the transcriptome, we also detected hundreds of unique barcodes per animal that connect these cells in their lineage (Additional Figure 2).

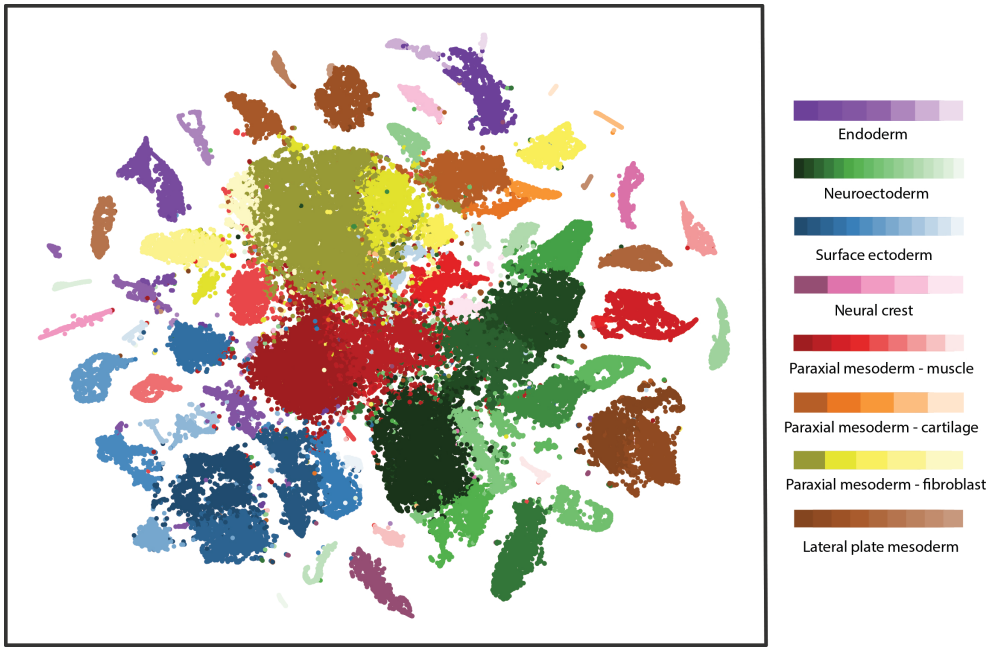


Figure 8. t-SNE representation of identified cell types for zebrafish larvae (5 dpf).

Cell types were grouped into 8 categories according to their known developmental origin, as indicated by the color code.

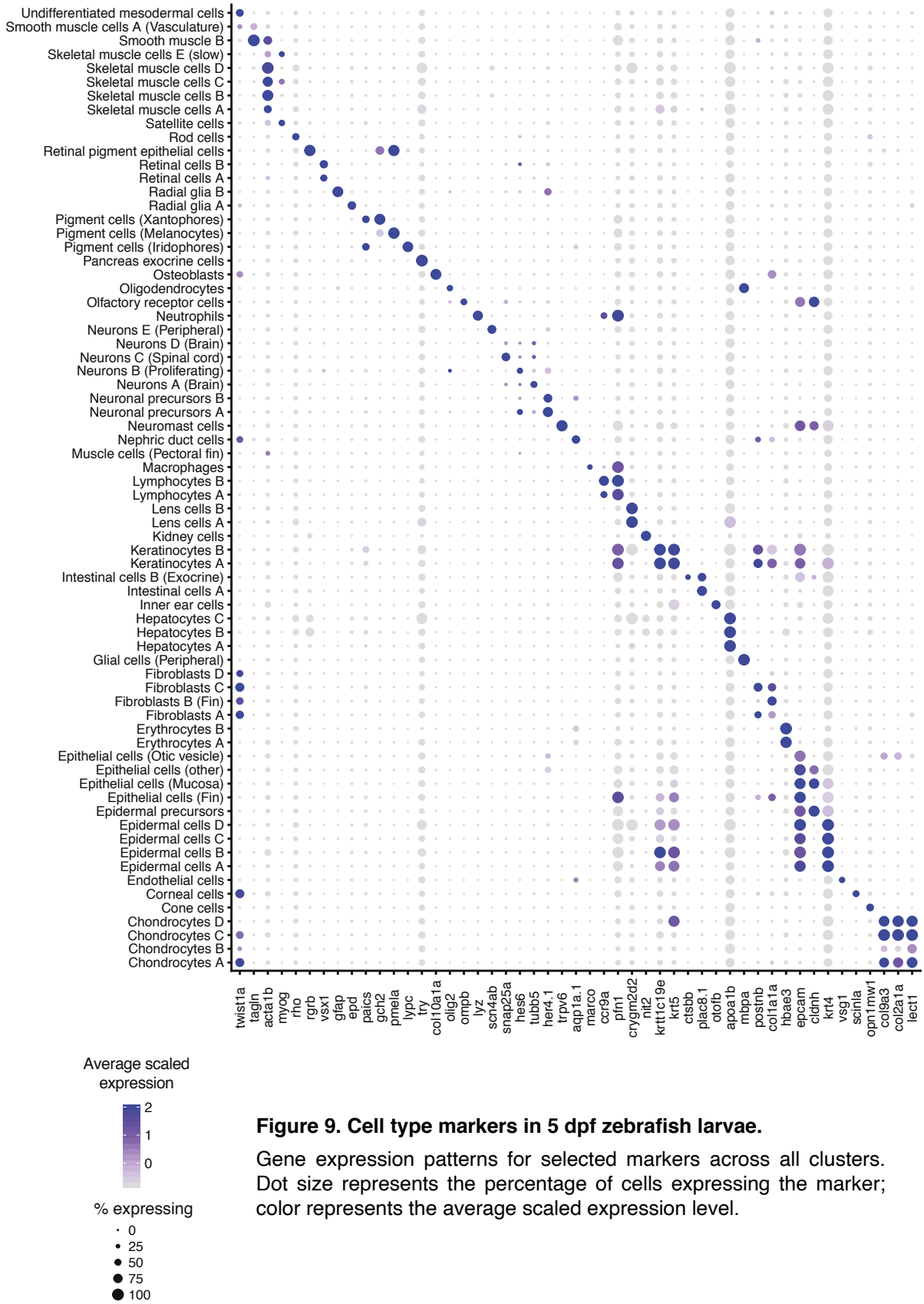


Figure 9. Cell type markers in 5 dpf zebrafish larvae.

Gene expression patterns for selected markers across all clusters. Dot size represents the percentage of cells expressing the marker; color represents the average scaled expression level.

Different cell types expressed barcodes with different efficiency

To quantify the barcode detection efficiencies in different cell types and from different genomic loci, we sequenced two offspring of Cas9 injected female fish - every cell in one offspring carried the same barcode profile it inherited from the fertilized oocyte.

We found that lineage barcode detection efficiency depended on cell type (Fig. 10), ranging from ~2 for erythrocytes to ~5 for epidermal cells. This observation probably reflected differences either in cell size or in the *ubi* promoter activity in that cell type.

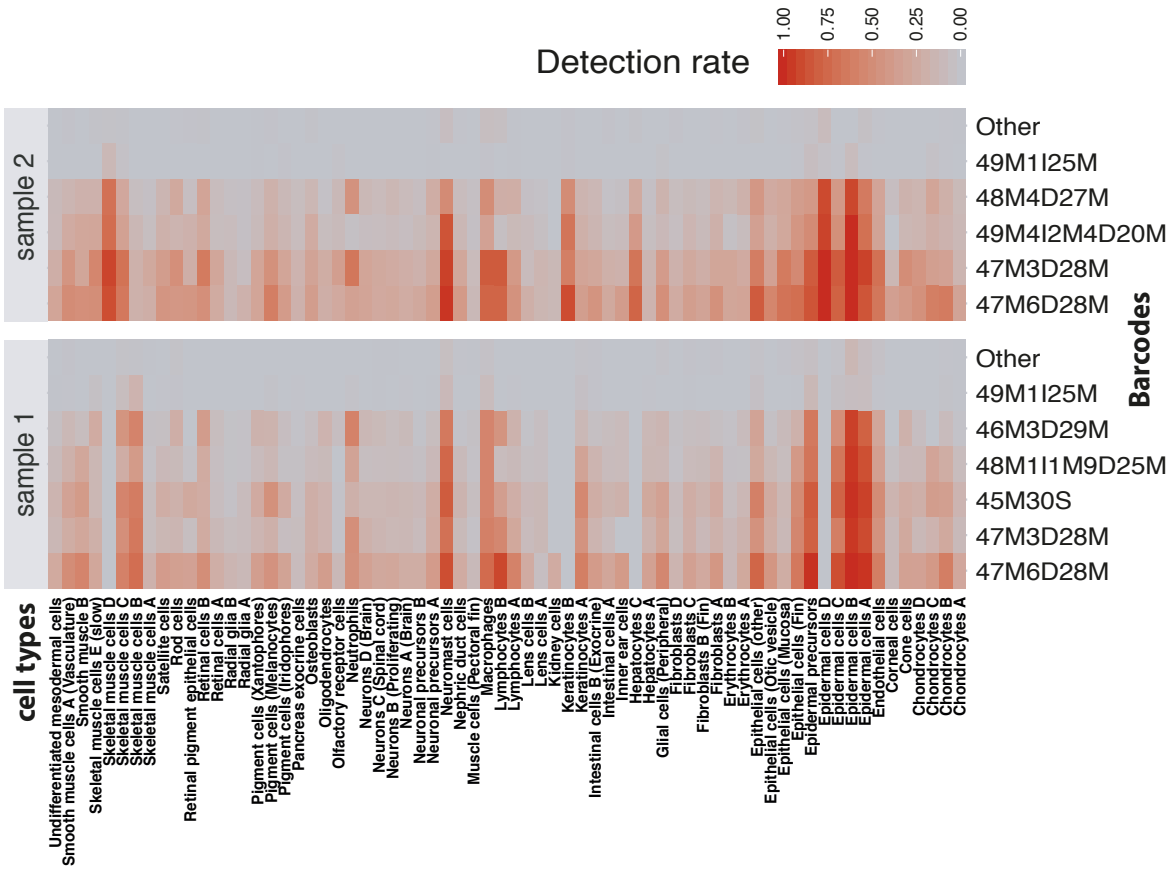


Figure 10. Single-cell lineage barcode detection rates.

The detection rate of barcode sequences (fraction of cells within one cell type where this barcode could be found), here represented by its CIGAR-string, are shown for all cell types. Non-wildtype sequences detected that have a very low detection rate are classified as sequencing errors and named “others”.

Importantly, this difference in expression efficiency between cell types was not barcode sequence specific - a cell type either expressed all barcodes highly or lowly. This ruled out selective silencing of a specific locus during development, which would be problematic for lineage reconstruction. However, we found one barcode with significantly lower expression. This kind of lowly expressed scars could be due to its genomic location (for example when the particular copy of the transgene is in a genomic location that is not very transcriptionally active). Such lowly expressed lineage barcodes carry little information, and could be potentially confused with sequencing errors (named “others” in Fig. 10). They will be removed from further analysis and tree reconstruction (see methods).

Lineage tree building

(The development of the lineage tree building algorithm and the later analysis of the trees were the work of Bastiaan Spanjaard, the co-author of the published paper. However, for the sake of completeness, the general principles and considerations of the lineage tree reconstruction are described here.)

We next aimed to use our barcode information to reconstruct a lineage tree of the cells from the scRNA-seq data. Ideally, we wanted to be able to reconstruct the barcode creation process based on the barcodes we sampled from the cells with scRNA-seq, which would result in a tree-like structure (Fig. 11a).

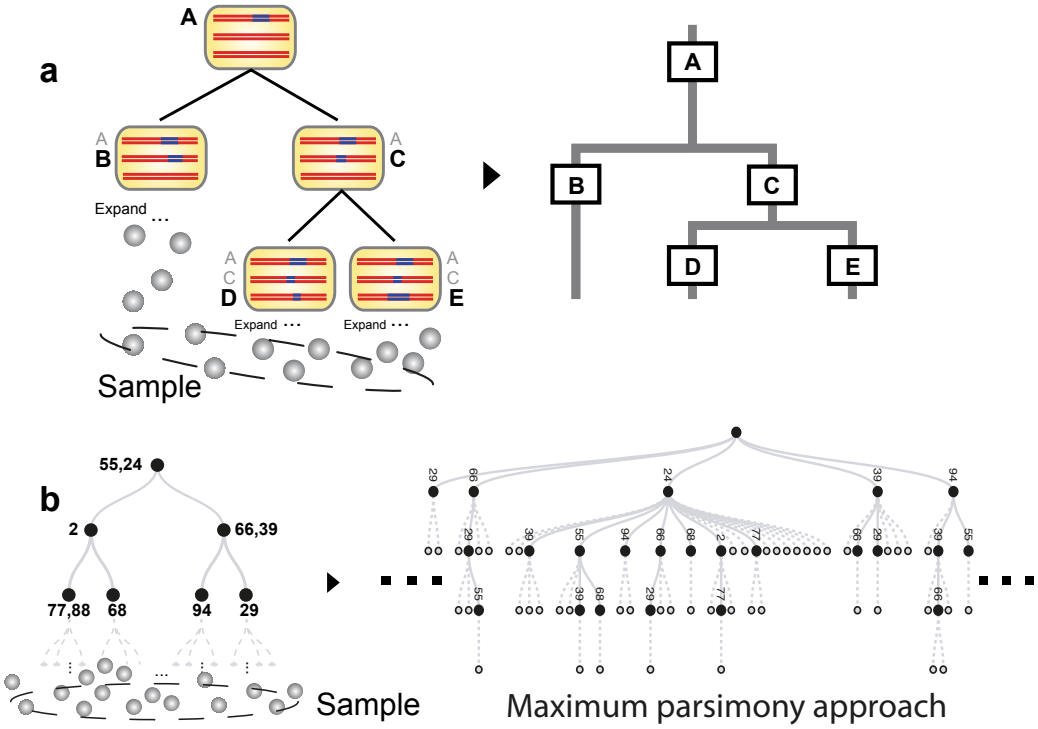


Figure 11. Maximum parsimony failed to reconstruct a lineage tree from our data.

a. A simplified barcode creation process with 5 barcodes created (symbolized by A-E). The newly created barcodes are shown in bold font. After the barcode creation ends, the cells will keep expanding and are sampled randomly at a later time point. The corresponding tree structure that ideally should be built from the sampling result is shown on the right side. **b.** Tree building with maximum parsimony and simulated data. The sampling here is simulated with dropout events. The parsimony algorithm failed to recreate the simulated lineage trees.

Methods like the maximum parsimony approach were used to reconstruct similar data before[40]. However, it was challenging to use such methods to reconstruct a correct lineage tree from our data. A major limitation here were the dropout events. A dropout event, which is inevitable in high-throughput scRNA-seq, occurs when the reverse

transcription failed to capture any transcript of a certain gene in a cell, mostly due to its low expression. In fact, a very simplified simulation with realistic dropout rates using the Camin-Sokal maximum parsimony resulted in a lineage tree that has much higher complexity than the true tree (Fig. 11b). Therefore, we needed to invent a new algorithm that is tailored to our data structure.

The new lineage tracing algorithm utilized the fact that a lineage tree following the barcode creation events can be converted into a network graph (Fig. 12, info-box): every barcode is a node in the network, and two nodes are connected with each other if the two barcodes are found at least once in the same cell. Since barcodes created

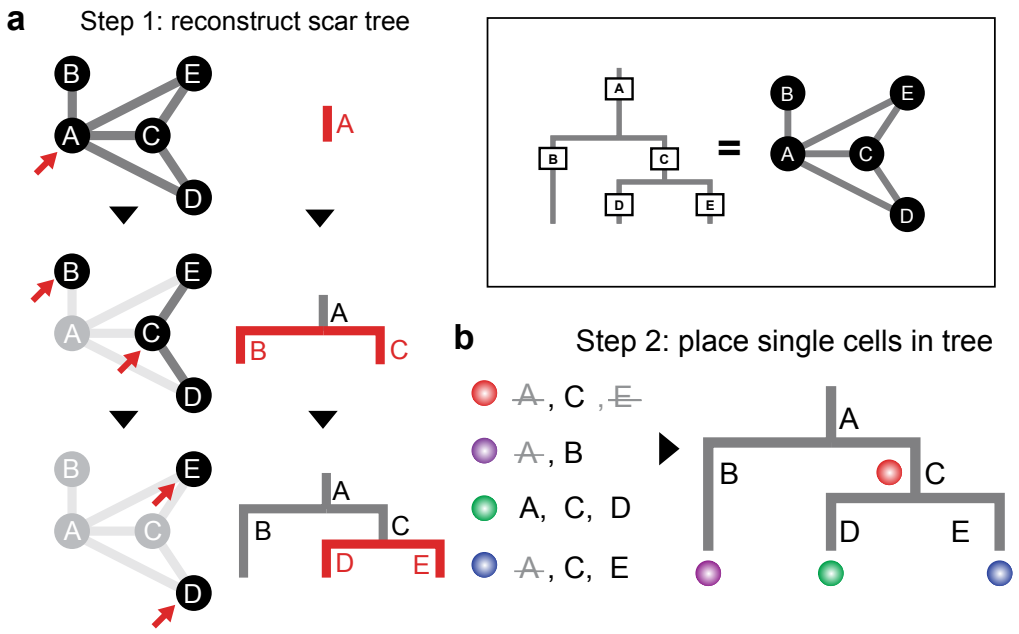


Figure 12. Graph based lineage tree reconstruction.

Info-box: A tree that is a reduced representation of the order of barcode creation events can be represented as a network graph. In a scar network graph, each node corresponds to a different scar, and scars that are co-expressed in the same cell are connected.

a. Iterative approach of tree reconstruction. The first barcode created is determined as the one with the highest connectivity (red arrow). Upon removal of the first barcode and its connections, the following barcodes are identified as the most highly connected ones in the reduced network.

b. After the barcode tree building, individual cells are placed in the tree according to their lineage barcode profile. Loss of barcode information in single cells (black letters: detected barcodes; gray crossed-out letters: undetected barcodes) leads to placements in the middle of the tree (for example, the red cell can only be placed at the second lowest branch of the tree due to the loss of information of barcode E). However, missing barcodes can also be inferred. (For example, the barcode A in the blue cell is inferred, since it expresses barcode C and E).

early should be found in more cells, the node with the most connections would be most likely to be the earliest scar that was created.

The lineage tree was then reconstructed on the basis of this network graph in a top-down iterative testing approach (Fig. 12a, see also methods). The cells were placed in the lineage tree in the corresponding branches afterwards. In this way, single barcode dropouts in particular cells were not relevant, as long as the “latest” barcode was detected in this cell. Hence, this method was more suited to deal with dropout events (Fig. 12b). The approach was tested for its superior performance towards the standard maximum parsimony approach (Additional Figure 3).

The reconstructed lineage tree of the 5dpf larvae showed a separation of cell types following known developmental principles like the split of ecto- and mesendodermal cells (Fig. 13a). The versatility of this kind of barcode-based tree building approach is shown when we “zoomed in”, only showing cells of the lateral plate mesoderm: a potential split of primitive erythrocytes can be observed within these lineages (Fig. 13b). We named our new approach of combined sequencing-based lineage detection and graph-based lineage tree reconstruction LINNAEUS (**lin**eage tracing by **n**uclease-**a**ctivated **e**diting of **u**biquitous **s**equences).

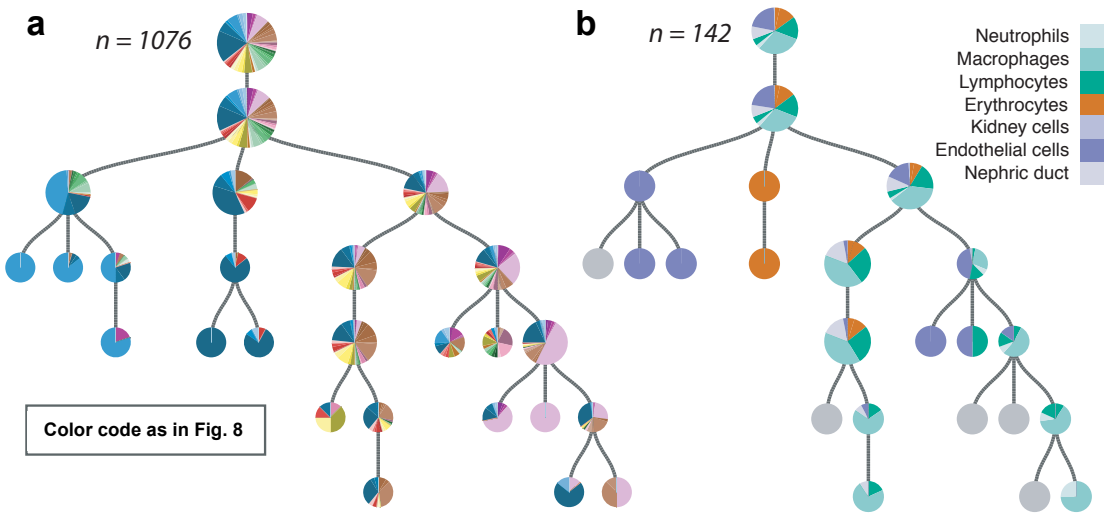


Figure 13. Example of a reconstructed lineage tree for a 5 dpf larva.

a. Lineage tree for 5 dpf larva. Pie charts are plotted small for n<50, medium for n≥50, and large for n≥1000. Color code for cell types as in Fig. 8. Lineage barcodes with p≥0.001 and barcodes that were detected in more than 1 larva were excluded from the analysis. In general, developmental lineages (ecto, meso, and endoderm) separate well in the tree. However, since barcoding ends at ~10 hours post fertilization, the end points of the branches may still give rise to multiple cell types in multiple tissues. **b.** Lineage tree for 5 dpf larva, zoomed into lateral plate mesoderm (see color code). The tree structure was determined based on the whole dataset.

Validation of LINNAEUS in adult organs

Next, we validated the applicability of our approach in the adult zebrafish. For this purpose, we sequenced single cells from four organs that are known to derive from the three developmental layers of the embryo: telencephalon from the ectoderm, heart from the mesoderm, pancreas and liver from the endoderm. After clustering and cell type analysis, we could find all expected major cell types in all three organs, including blood and immune cells that are sequenced along with these tissues (Fig. 14 and Fig. 15 for overview of the clustering, for detailed t-SNE and gene lists see also Additional Figure 4 and Additional table 3).

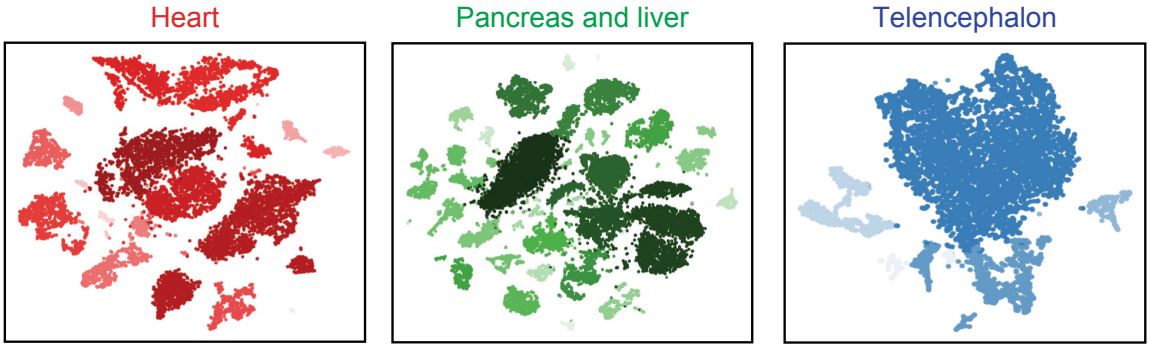


Figure 14. t-SNE representation of cell types in different adult zebrafish organs.

t-SNE representations of scRNA-seq data for dissociated organs from adult zebrafish (*n* = 3 animals). Cell types of the same organ are shown in shades of the same color for better distinction later in the lineage trees.

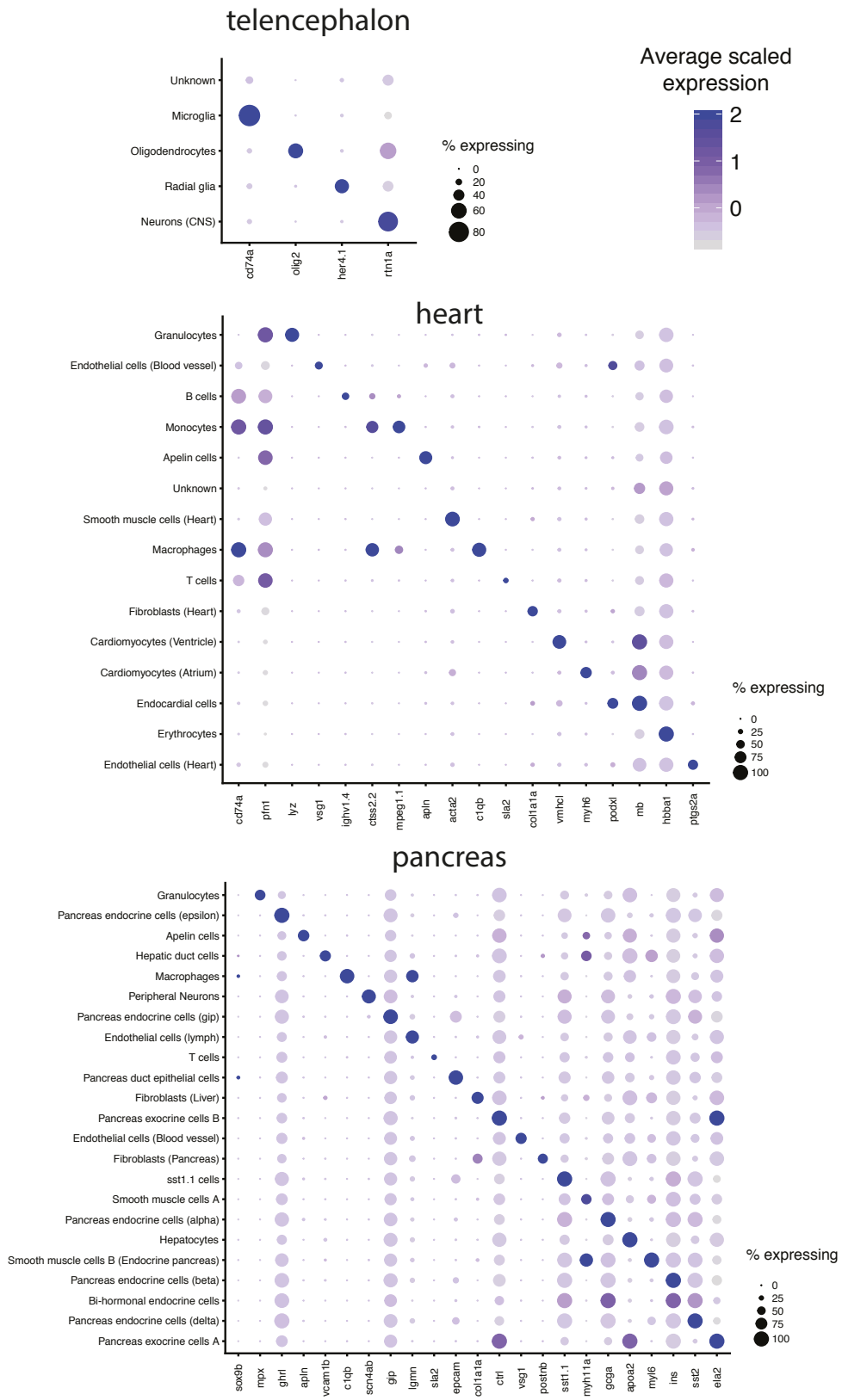


Figure 15. Marker genes for cell type identification in the adult zebrafish organs.

Gene expression patterns for selected markers across all cell types in dissociated heart, pancreas and telecephalon. Dot size represents the percentage of cells expressing the marker; color represents the average scaled expression level.

After lineage tree reconstruction, we first found that immune cells from all organs co-localize in the lineage tree (Fig. 16a), even though they were sequenced in different libraries per organ. This served as an additional validation for the approach, since our algorithm correctly determined the same lineage of immune cell types in all different adult organs.

Furthermore, we saw a lineage split in our lineage tree between cells from the endodermal organs (pancreas, liver) and the mesodermal organ (heart) (Fig. 16b). The brain cells were underrepresented in the trees because of their low expression of the lineage barcodes: the barcodes were only detectable in a few cells of the brain.

We also confirmed that we could distinguish lineage differences between different cell types in the same organ, e.g. the lineage of endocardial cells from the fibroblasts in the heart, or the lineage of the exocrine cells in the liver form the hepatocytes (Additional Figure 5). This made us confident that we could use this approach to study the zebrafish heart regeneration.

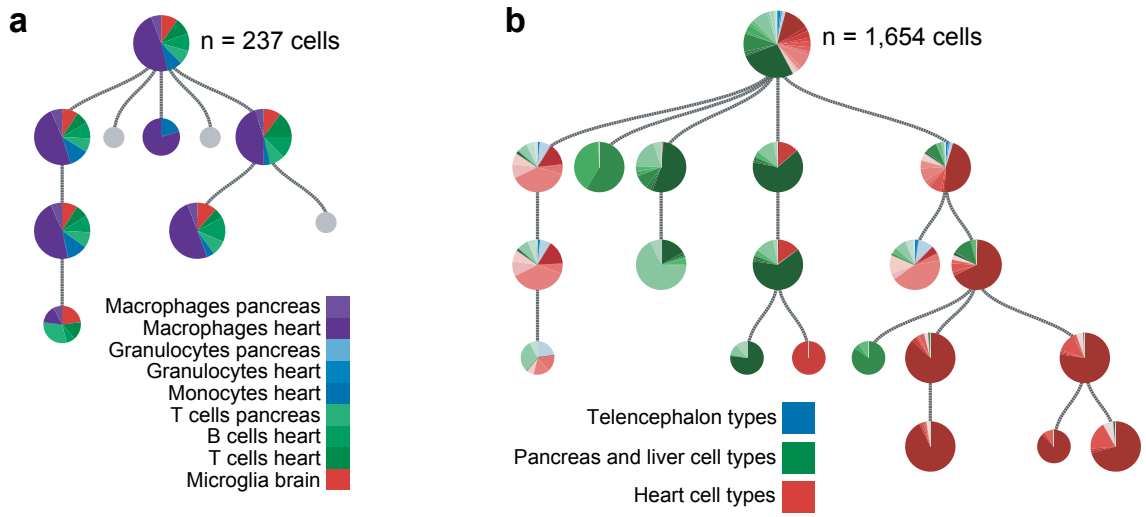


Figure 16. Lineage tree for organs from one adult.

a. Lineage tree zoomed into immune cell types (see color code). As expected, immune cells from different organs cluster together in the lineage tree. **b.** Pie charts are plotted small for $n <$ 50, medium for $n \geq 50$, and large for $n \geq$ 1,000. Scars with creation probability $\geq$ 0.01 were excluded from the analysis. Color code as in Fig. 14.

Single-cell RNA-seq of the regenerating zebrafish heart

After we successfully established LINNAEUS in larvae and adult organs, we set out to implement this method in the study of the zebrafish heart regeneration.

Healthy and injured zebrafish hearts were collected at different stages of regeneration (see experimental strategy and methods). The approach described above was used for sequencing, mapping, filtering, and clustering. We were able to observe all cell types that we had been expecting in our single-cell sequencing data (For detailed clustering and differentially expressed genes per cluster see Additional Figure 6 and Additional table 4). In line with the existing literature, we observed an increased number of immune cells and fibroblasts at 3 or 7 dpi, respectively (Fig. 17).

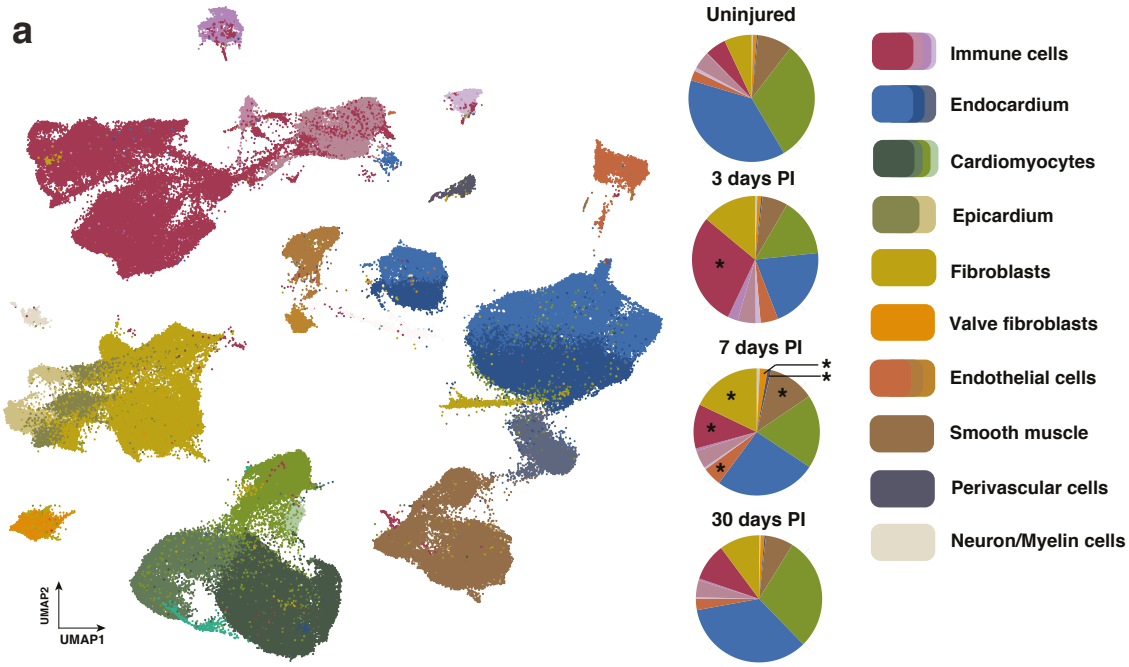


Figure 17. UMAP representation of scRNA-seq data and clustering results.

Pie charts show the proportions of different cell types at different time points after injury. In the pie chart representation, similar cell types are grouped and shown by one (representative) color. Asterisks (*) denote cell types with a statistically significant change in proportions compared to uninjured control. Similar cell types (e.g. different populations of immune cells) are shown in similar colors. For detailed color scheme see also Additional Figure 6.

Cell types have spatially distinct gene expression profiles

Interestingly, the cells of all three layers of the heart (epicardium, myocardium, and endocardium) showed a substantial amount of heterogeneity in their gene expression that is not related to the injury state (Fig. 18).

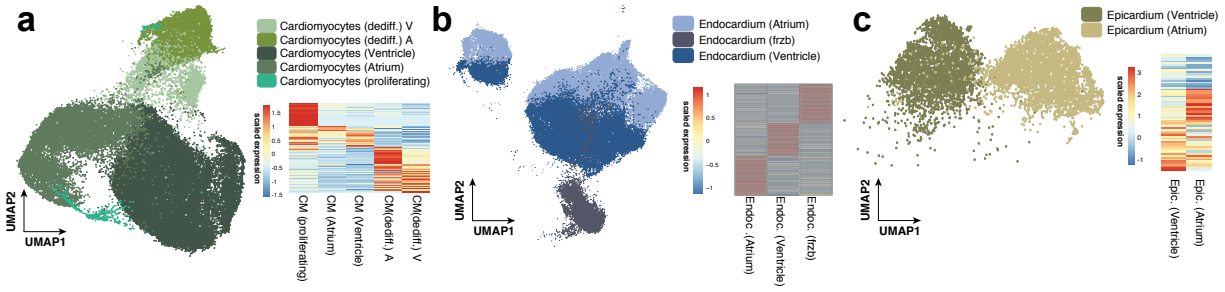


Figure 18. Sub-clustering of cardiomyocytes (a), endocardium (b), and epicardium (c).

Heatmaps show top differentially expressed genes of each sub-type. The genes used for the heatmap can be found in Additional table 4. The UMAP representations of the clustering are zoomed in from Fig. 17.

We hypothesized that the spatial localization in the heart could be causing this heterogeneity of these subtypes, which has been described for the cardiomyocytes before[21]. By mapping our single-cell data onto a spatially-resolved bulk dataset (tomo-seq[41]) from a healthy zebrafish heart (scRNA-seq deconvolution[42], see methods), we could strongly suspect the atrial vs. ventricular localization of these cells (Fig. 19a). We could confirm this by physically separating the two cardiac chambers prior to scRNA-seq (Fig. 19b).

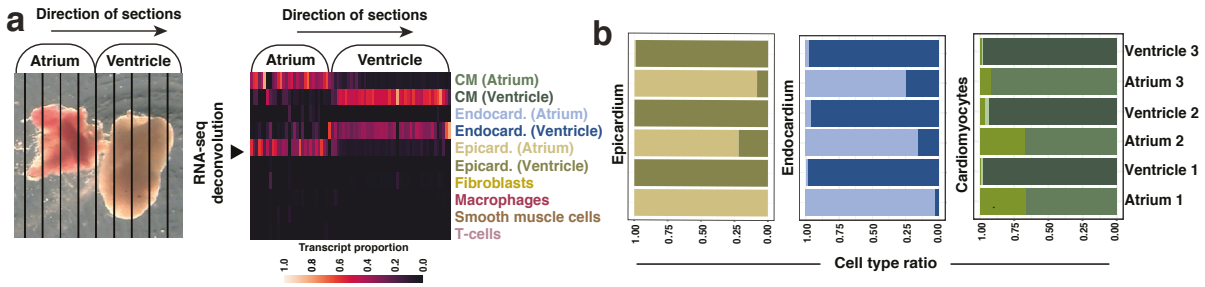


Figure 19. Localization of sub-cell-types.

a. Mapping of single-cell data onto a spatially-resolved tomo-seq dataset. A computational deconvolution approach reveals chamber-specific sub-cell-types. The direction of the tomo-seq sections are shown. The result of the mapping is shown as a heatmap. Most of the subtypes with spatial differentiation could be mapped either to the atrial, or the ventricular sections. **b.** Distribution of subtypes of cardiomyocytes, endocardial cells, and epicardial cells for three scRNA-seq datasets in which atrium and ventricle were physically separated. Color scheme as in Fig. 18.

In addition to the spatial differences in gene expression, we also found a sub-population of cardiomyocytes (CMs) that increased in numbers at 3dpi (Fig. 20a). Those, unlike constitutive adult CMs, which differentially express genes involved in ATP synthesis and the tricarboxylic acid cycle (*atp5pd*, *aldoaa*), express genes involved in cardiomyocyte development (*ttn*, *bves*, *synpo2lb,* Fig. 20b).

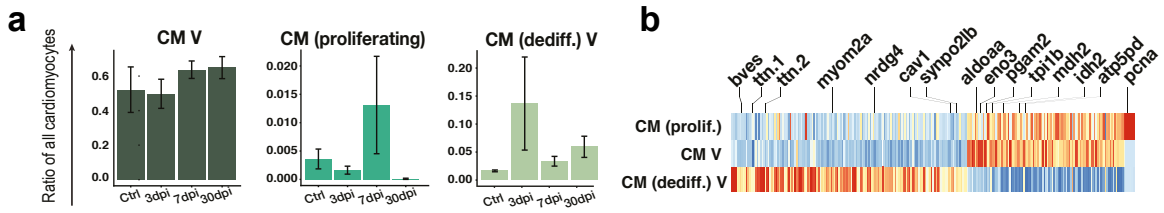


Figure 20. a. Relative changes of abundance for different subtypes of cardiomyocytes (CM) across the timepoints (error bars show standard error of the mean). **b.** Differentially expressed genes between subtypes of cardiomyocytes.

Fibroblasts in the regenerating zebrafish heart

We next focused our attention on the cardiac fibroblast population, the main contributors to the heart fibrosis after injury. In these cells, we were able to observe the expressions of important factors that are known to positively regulate regeneration[43,51,52], e.g. aldh1a2, which is involved in retinoid acid (RA) signaling[51] (Fig. 21a). The sub-clustering of all fibroblasts (here defined as cells expressing *col1a1a*, a bona-fide fibroblast marker) showed many subtypes among them (Fig. 21b, Additional table 5). They expressed different factors related to the extracellular matrix (ECM) (Fig. 21c), indicating differences in their function.

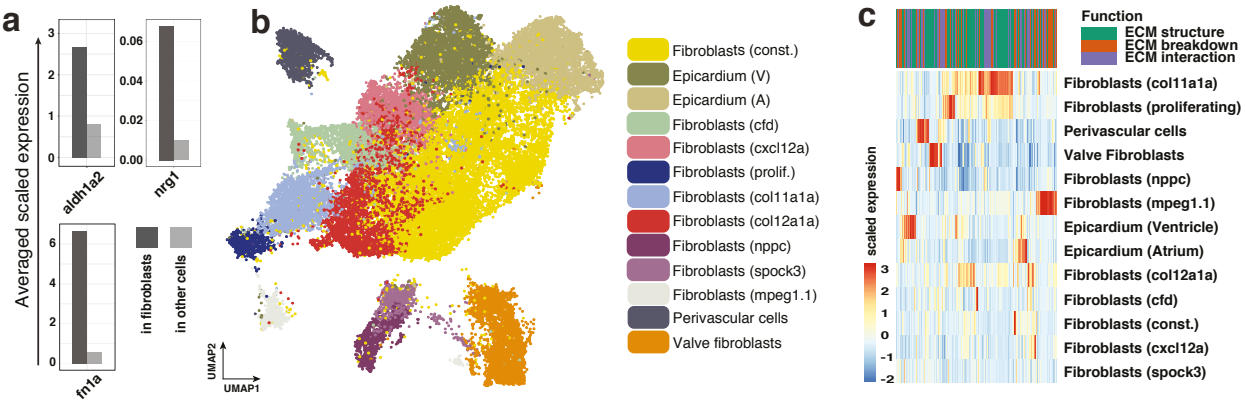


Figure 21. Cell type diversity of cardiac fibroblasts.

a. Comparison of average expressions of known pro-regenerative factors in fibroblasts and other cell types. **b.** UMAP representation of the subclustering of *col1a1a* expressing cells. **c.** Expression of extracellular matrix (ECM) related genes in different fibroblast cell types. The genes are classified according to their contribution to structure, breakdown or interaction of the ECM.

A time-resolved analysis of the abundance of different fibroblast subtypes showed, in addition to a proliferative population, three other transient fibroblast subtypes (Fig. 22): *col11a1a* positive fibroblasts, *col12a1a* positive fibroblasts, and *nppc* positive fibroblasts. Some other cell types also slightly increase during regeneration, like the perivascular cells and the valve fibroblasts.

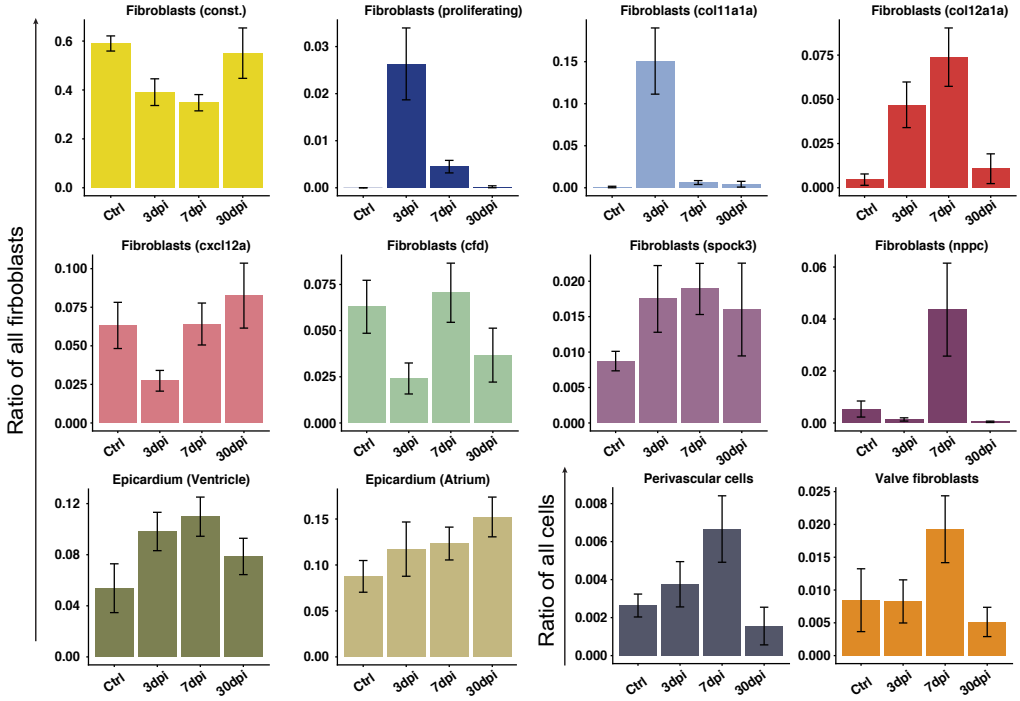


Figure 22. Cell type dynamics of cardiac fibroblasts.

Cell number dynamics of all fibroblast subtypes during regeneration. Mean value across all replicates of the same time point is shown, error bars indicate standard error of the mean. Cell type names indicate a differentially expressed gene characteristic for a cell type. Color scheme is the same as in Fig. 21b.

A thorough study of the gene expression of these transient subtypes showed that many of them expressed different pro-regenerative factors (Fig. 23a, a heatmap including all cell types is shown in Additional Figure 7). For example, *nrg1*, previously shown as a CM mitogen[43], was expressed highly in *col12a1a* fibroblasts, and *aldh1a2*, the enzyme that is responsible for retinoic acid (RA) production, was expressed highly in the epicardial cells and *nppc* fibroblasts. Interestingly, *stra6*, a gene that responds to RA-signaling, was only expressed in proliferating, *col11a1a*- and *col12a1a* fibroblasts, but not in other cell types.

The localization of these transient cell types, shown by fluorescent in-situ hybridization (Fig. 23b, more replicates see Additional Figure 8), indicated that they are all localized in approximation to the injury area, which strongly suggested their contribution to a regenerative niche, in which they would create a physical and chemical environment beneficial for regeneration.

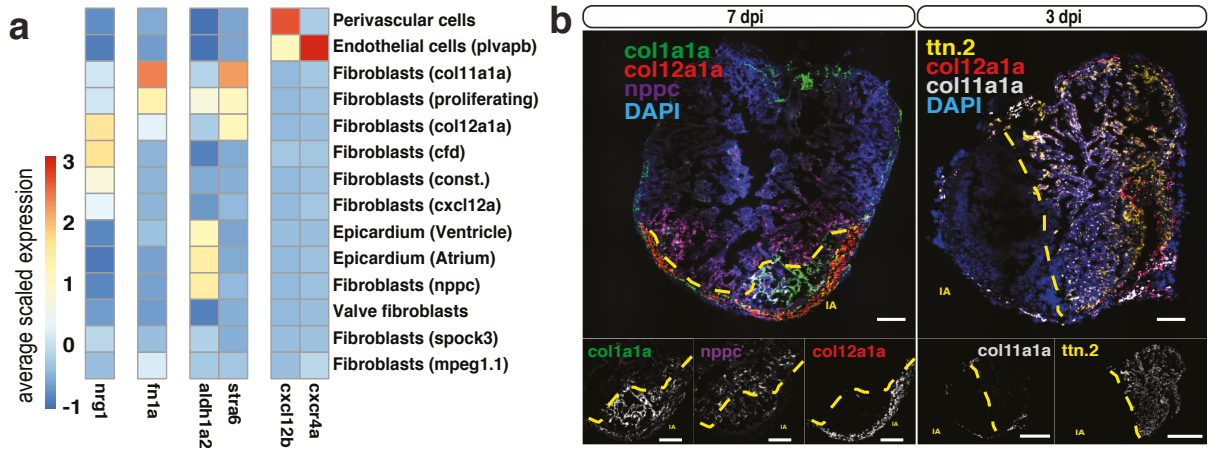


Figure 23. a. Average expression of selected signaling genes in fibroblast sub-clusters. Blood vessel endothelial cells were added to show their interaction with perivascular cells via cxcl12b-cxcr4a signaling. **b**. Fluorescent in-situ hybridization of marker genes. Left panel: Fibroblasts (const.) (green), Fibroblasts (*col12a1a*) (red), Fibroblasts (*nppc*) (purple) at 7 dpi. Right panel: Fibroblasts (*col12a1a*) (red), Fibroblasts (*col11a1a*) (white) and dedifferentiated cardiomyocytes expressing *ttn.2* (yellow) at 3 dpi. Injury areas are indicated with dotted yellow line. Scale bar: 100µm.

Origin of transient cell types

To understand the temporal process of their activation, we next intended to determine the origin of these transient fibroblasts. With our LINNAEUS approach, we were able to build lineage trees of the regenerating heart at different time points after injury. Due to the nature of our lineage tracing method, each biological replicate we analyzed resulted in a separate lineage tree (Fig. 24). While these trees showed common characteristics as described before (separation of immune cells, separation of endocardium and fibroblasts), they were also very distinct from each other in their tree topology due to the stochastic nature of the barcode creation process. (Trees from all samples are shown in Additional Figure 9).

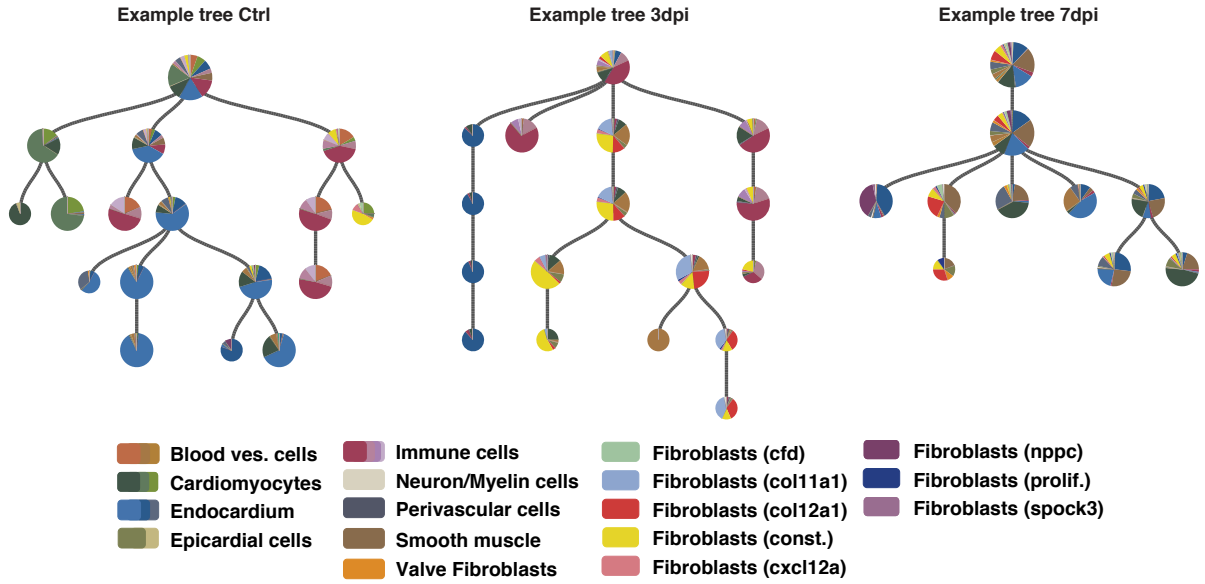


Figure 24. Examples of lineage trees from different time points after injury.
One example tree each are show for the healthy heart, an injured heart at 3dpi and an injured heart at 7dpi. Cell types are indicated by different colors. The color scheme is the same as in Fig. 17a and Fig. 21b.

For a systematic analysis of lineage relations using all lineage trees, we needed to find a method that allows us to integrate lineage data from multiple samples. We rationalized that in our lineage trees, a node indicated a common developmental ancestor for those cells within, and any transient cell type would be found in the same node like its progenitor. Ergo, for cell types which were related by lineage, a strong correlation between cell type ratios across all single nodes should be found (Fig. 25.

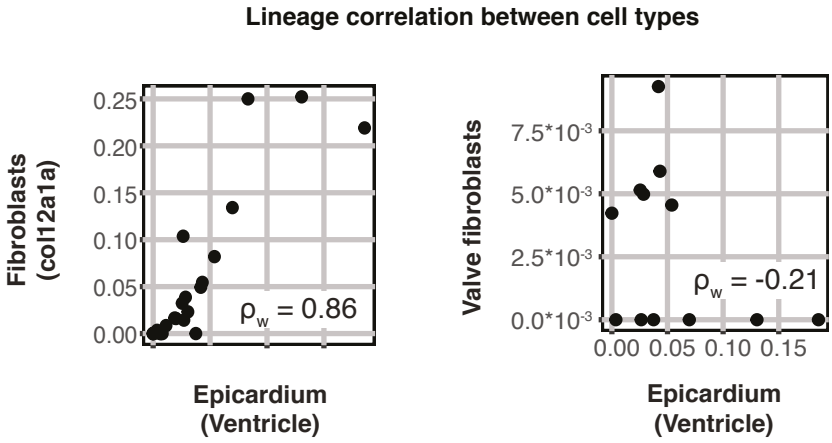


Figure 25. Examples of weighted correlations between cell type ratios in single nodes.

We calculated weighted correlations of cell types in single nodes over several trees in order to quantify lineage similarity. Two examples are shown here for either high (left) or low (right) correlation between two cell types.

see methods for details). In this way of node-focused analysis, we would circumvent the tree topology, and even use the stochasticity in our data to our advantage.

We determined the fractions of all cell types for all nodes (Additional Figure 10a and 10b) for 3 and 7 days after injury, and calculated their correlations. Next, we performed hierarchical clustering of these correlations for both time points (Fig. 26). All immune cells were found in the same cluster, indicating their common lineage, which validated our approach. At 3 dpi, the earliest timepoint at which we detected the transient *col11a1a* positive and *col12a1a* positive fibroblasts, we found these two cell types in the same cluster as constitutive fibroblasts and the epicardial cells. The same

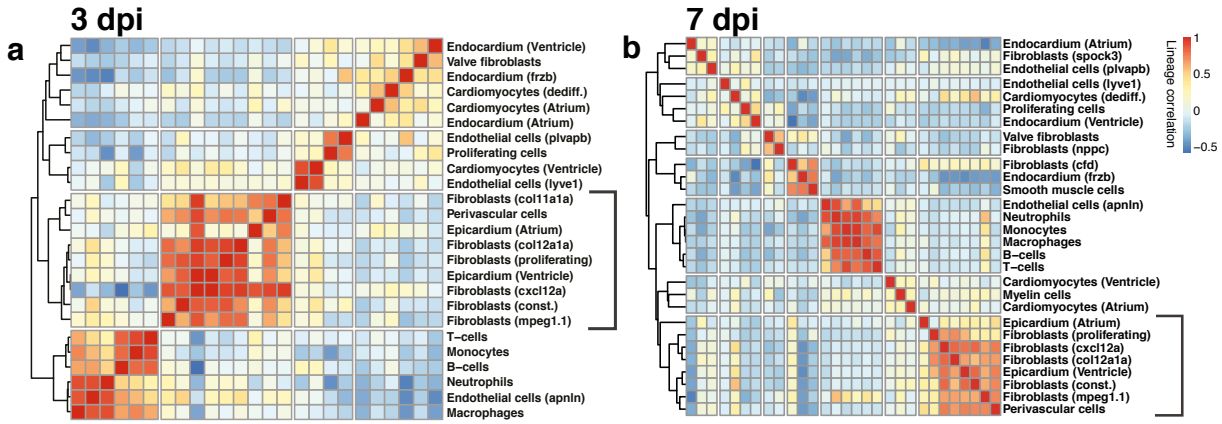


Figure 26. Clustering of cell type lineage correlations.

Clustering of pairwise lineage correlations of all cell types at either 3 dpi **(a)** or 7dpi **(b)**. At both time points, a stable cluster of many fibroblast cell types was found (brackets), this cluster included also the epicardial cells. Additionally, a cluster of immune cells was found at both time points.

was observed for them at 7 dpi, the earliest time point at which we detected the transient *nppc* fibroblasts. Additionally, a downsampling of 50% for the cell number in the trees confirmed the stability of this clustering and our observation. Hence, our data strongly suggested that *col11a1a* positive fibroblasts and *col12a1a* positive fibroblasts derive from the epicardial lineage.

This hypothesis was confirmed by a genetic lineage tracing experiment using the Cre-Lox transgenic line *Tg(tcf21:Cre^{ERt2}; ubi:Switch)*[44,45] (see methods). The Tcf21 promoter was previously shown to be active in epicardial derived cells[44]. After the induction of Cre, the expression of Tcf21 driven mCherry indeed co-localized with the expression of *col12a1a* (Fig. 27a). We further elucidated short term cell type dynamics within the epicardial derived fibroblast subtypes by RNA velocity[46,47], which uses the intron-extron ratio of genes in single cells to infer the direction of transcriptome changes (see methods). The result of RNA velocity indicated strongly that *col11a1a* fibroblasts derive from constitutive fibroblasts, and *col12a1a* fibroblasts either derive from constitutive fibroblasts or the epicardium (Fig. 27b).

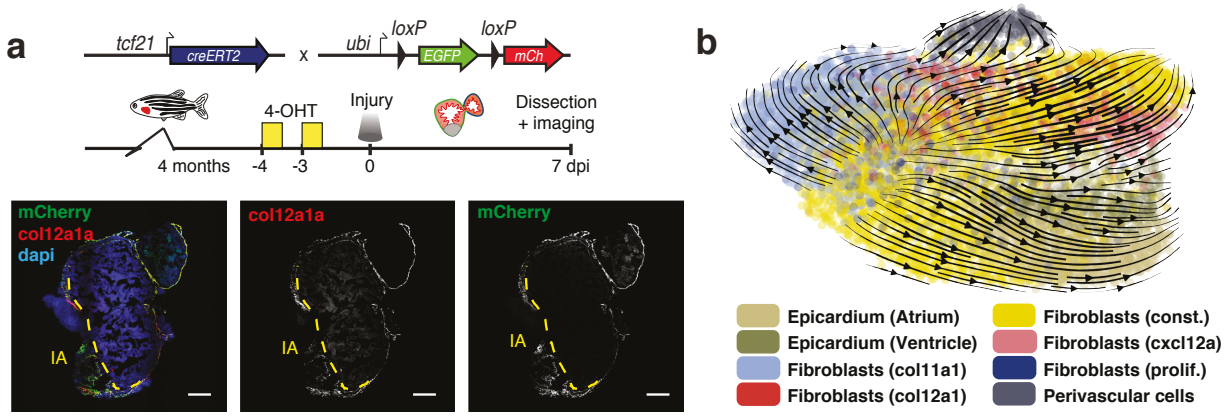


Figure 27. Epicardial origin of *col11a1a* /*col12a1a* fibroblasts.

a. Microscopy-based lineage tracing shows co-localization of tcf21:mCherry, an epicardial marker, and *col12a1a* expression, which marks both the *col11a1a*- and *col12a1a* fibroblasts. Scale bar: 100µm. **b.** Trajectory (RNA velocity) analysis suggested constitutive fibroblasts as the source of *col11a1a*- and *col12a1a* fibroblasts at 3 dpi. The color scheme is the same as in Fig. 21b.

Endocardial derived fibroblasts

Surprisingly, LINNEAUS suggested that several fibroblast cell types (*nppc*- and *spock3* fibroblasts, valve fibroblasts) were not part of the epicardial derived cluster, indicating that these fibroblasts may have a different origin. After further analysis of the lineage tree and the tree node correlations, we found the valve fibroblasts clustering with endocardial cells, and over 80% of the nodes containing *nppc* fibroblasts also contained different endocardial cell types (endocardium (Atrium), endocardium (Ventricle) and endocardium (frzb)). This suggested that they could be endocardium derived fibroblasts. And indeed, RNA-velocity confirmed the transition of ventricular endocardial cells to *nppc* fibroblasts (Fig. 28).

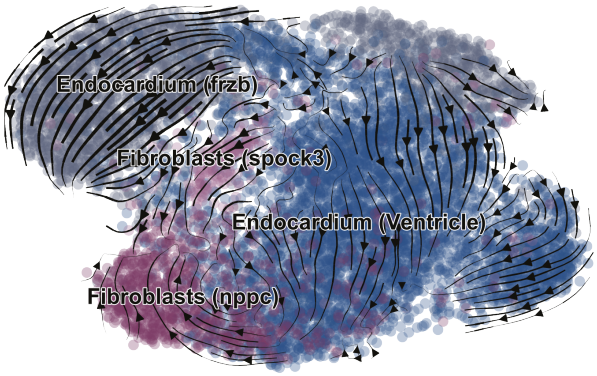


Figure 28. Endocardial origin of *nppc* fibroblasts.

Trajectory analysis indicates transitions from endocardial cells to *spock3*- and *nppc* fibroblasts. The color scheme is the same as in Fig. 17a and Fig. 21b.

Endocardium responds to injury by turning on a fibrotic program

We observed that *nppc* fibroblasts continued to express endothelial genes (e.g. *vwf, fli1a*) in addition to ECM genes, indicating that, to some degree, they maintained their endocardial gene expression after activating a fibrotic gene expression program. We rationalized that the appearance of these cells was an injury-specific response of the endocardium and would only take place when the endocardium was affected by the injury. To investigate this, we varied the time of contact of the cryo-probe with the heart. Considering the internal location of the endocardium, a very short contact should lead to less damage to the endocardium. We could confirm that longer injury time led to deeper injuries, measured by the injury area (Fig. 29a), and also led to much stronger induction of *nppc* expression in the deeper area of the injury (Fig. 29b). This very local nature of the endocardial reaction to injury could have also been the reason why a correlation analysis of lineage relationships, like described above, was not suited to find the precursors to *nppc* fibroblasts.

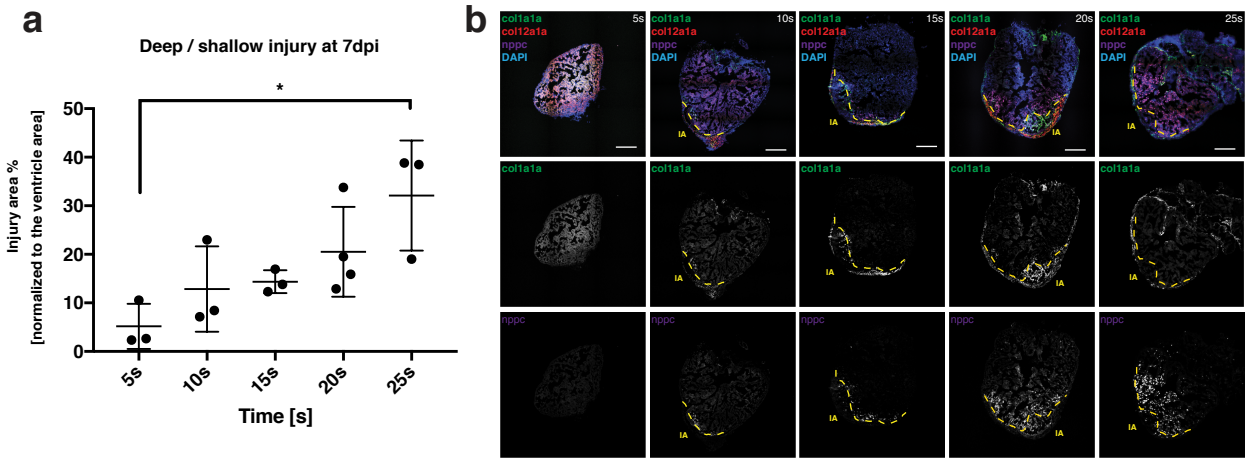


Figure 29. Deeper injuries cause higher *nppc* expression.

a. Injury areas (IA) in % were normalized to the ventricular area. n(5s)=3, n(10s)=3, n(15s)=3, n(20s)=4, n(25s)=3. Mean ± SD. *One-way ANOVA with Tukey's multiple comparison test. *P= 0.0128. **b.** Injuries of 20s and 25s penetrate the heart and generate *nppc* expression (purple) beyond the injury area. Scale bar: 100 μm.

The regeneration process is impaired after Wnt inhibition

Finally, with the advanced knowledge of cell types involved in the zebrafish heart regeneration process, we wanted to investigate which signaling pathway is required to generate these transient cells. Furthermore, with their lineage information, we can determine whether cells with a shared lineage origin exhibit similar changes after a stimulus, e.g. in perturbation experiments.

We noticed that Wnt signaling related factors are expressed in almost every cell type, except for immune cells, and are especially high within the fibroblasts (Fig. 30, Additional Figure 11).

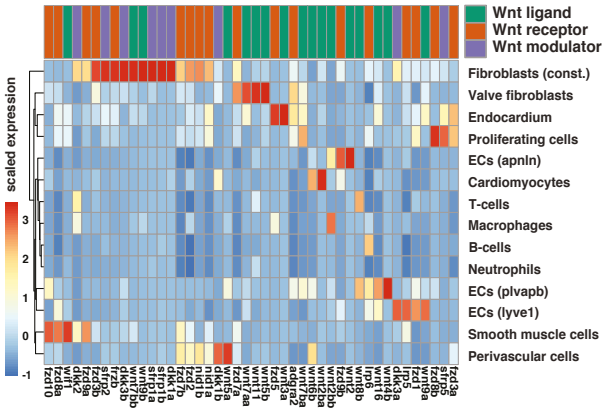


Figure 30. Expression of Wnt signaling factors across cell types of the zebrafish heart.

The selection of genes show all common canonical Wnt signaling factors.

We designed a Wnt inhibition experiment using the canonical Wnt-antagonist IWR-1[48,49] (see experimental strategy and methods) to dissect the effect of the absence of canonical Wnt signaling. After inhibition, we observed significantly delayed regeneration, while the injury and fibrosis were still present at 30dpi (Fig. 31, Fig. 32).

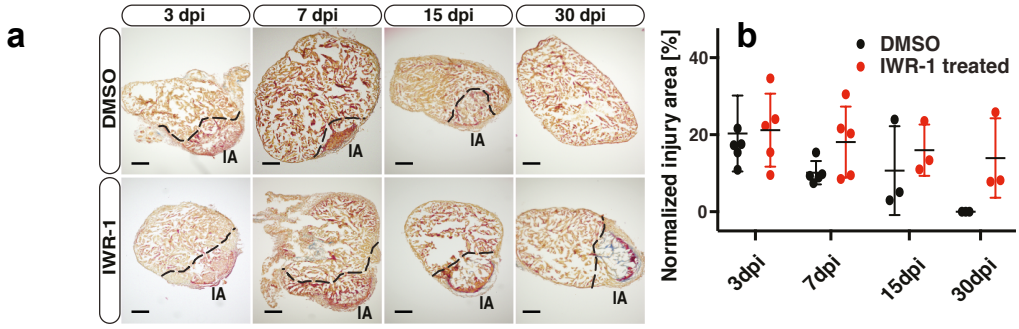


Figure 31. Wnt inhibition impaired the regeneration of the zebrafish heart.

a. Histological comparison of the injury area (IA) at different time points after injury with or without Wnt inhibition. The collagen depositions are stained here as blue. Scale bar: 300μm. **b.** Relative size of the injury area to the size of the heart across all histological replicates, mean and standard deviation are shown.

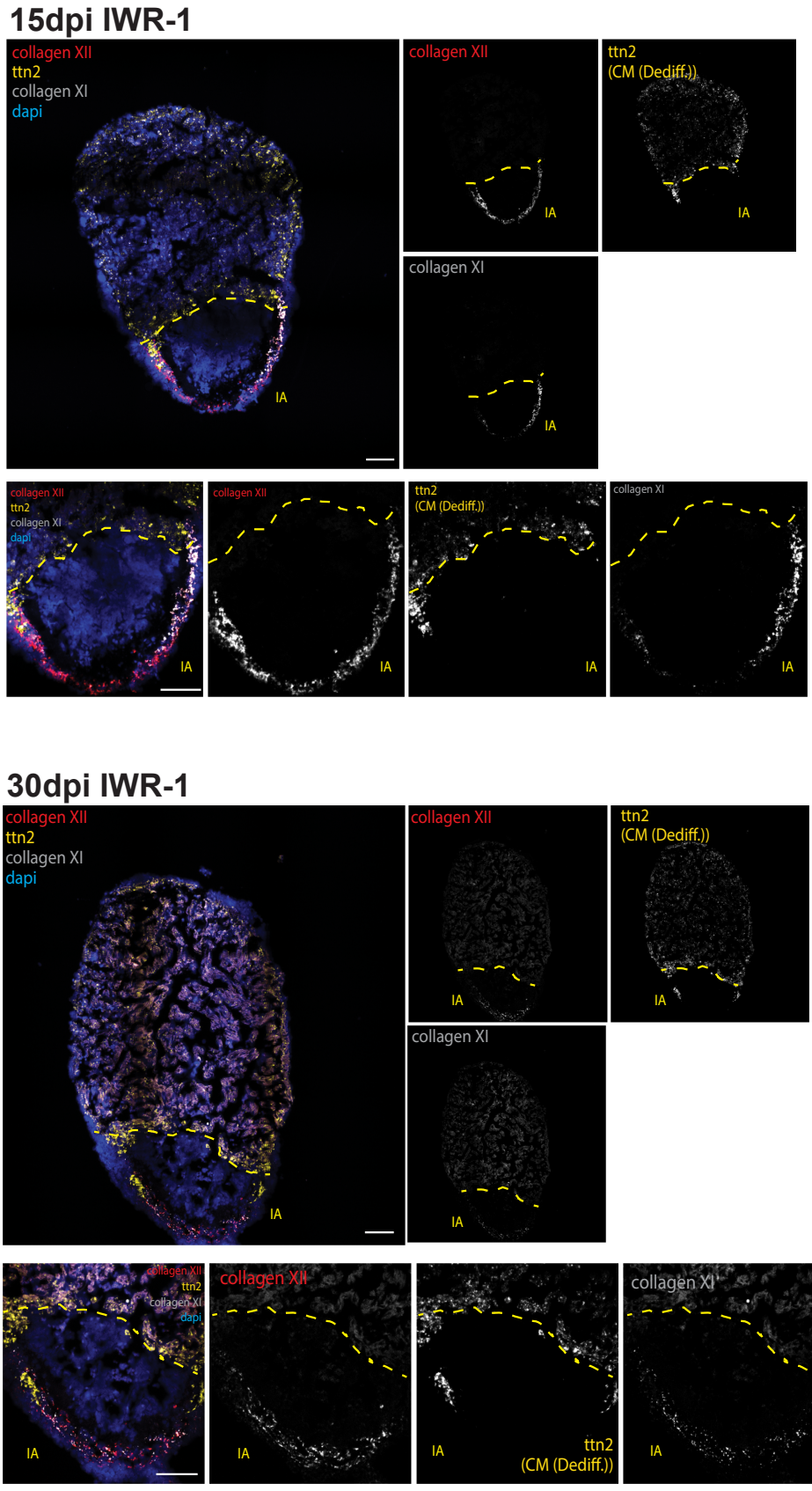


Figure 32. Fluorescent in situ hybridization at 15 and 30 dpi after Wnt inhibition.

The wound areas still exist at late time points after injury. Cell types present at early time points of injury are still present here.

Deregulation of dedifferentiated cardiomyocytes after Wnt inhibition

In addition to the macroscopic observation of impaired regeneration, we next aimed to describe the effect of Wnt inhibition in the regeneration process on a cellular level. Considering that almost every cell type expressed more than one factor involved in Wnt signaling (Additional Figure 11), it was remarkable that only several cell types changed in their abundance after IWR-treatment compared to DMSO control (Fig. 33).

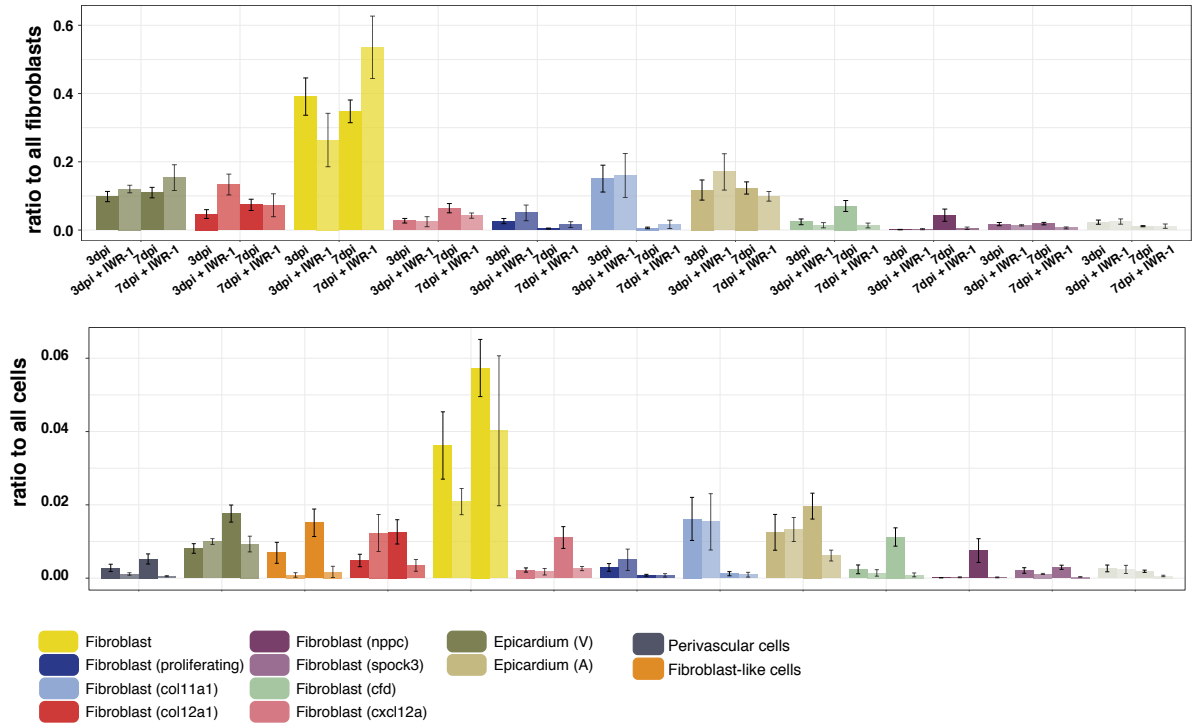


Figure 33. Cell type abundance changes after Wnt inhibition at 3 or 7 dpi.

The cell numbers here are either normalized to the number of all fibroblasts (upper panel), or to the number of all cells in the respective heart (lower panel). The color scheme is the same as in Fig. 21b.

First, we found that the previously described cardiomyocyte population, which is characterized by genes that are involved in cardiomyocyte development, is present in less abundance at 3dpi (Fig. 34a). At 7dpi, in-situ-hybridization showed an alteration in the spatial localization of these cells. While they were localized at the injury area at 7dpi in DMSO control samples, they were located in the border zone adjacent to the injury after Wnt-inhibition (Fig. 34b).

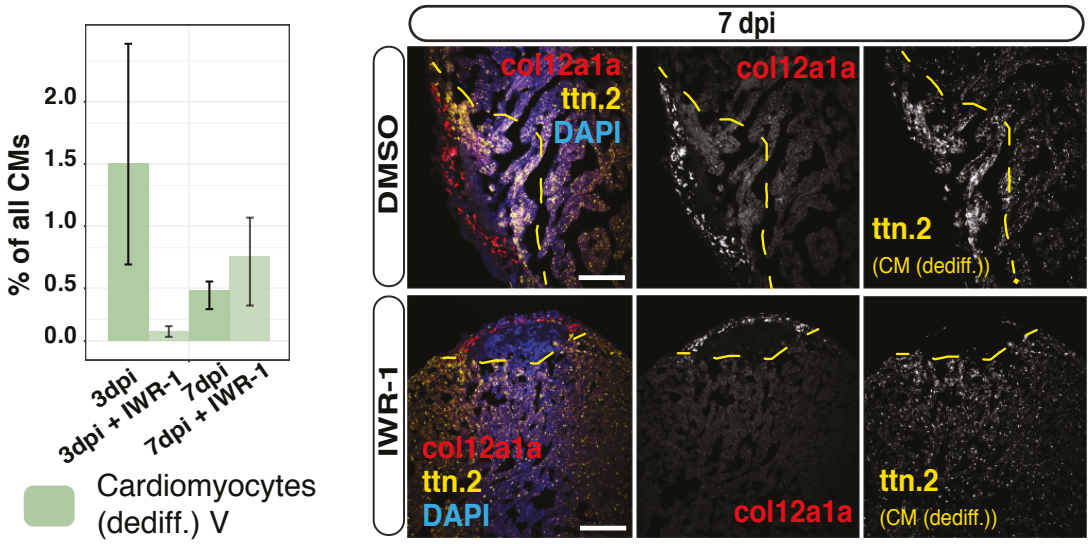


Figure 34. Dedifferentiated cardiomyocytes after Wnt inhibition

a. Abundance of dedifferentiated cardiomyocytes at 3 and 7 dpi with and without Wnt inhibition. Error bars show the standard error of the mean. (CM=Cardiomyocytes). **b.** Localization of dedifferentiated (*ttn.2*) cardiomyocytes (yellow) at 7 dpi in the injured heart with or without Wnt-inhibition. Yellow dotted lines indicate the injury area.

Endocardium derived fibroblasts diminished after Wnt inhibition

Intriguingly, we noticed that all endocardial derived fibroblasts, as identified by our lineage analysis (*nppc* fibroblasts, *spock3* fibroblasts, and valve fibroblasts) decreased in abundance after Wnt inhibition (Fig. 35). This strongly suggested that their activation is Wnt dependent. Most prominently, *nppc* fibroblasts, which appear only transiently after injury, were almost completely absent after Wnt inhibition.

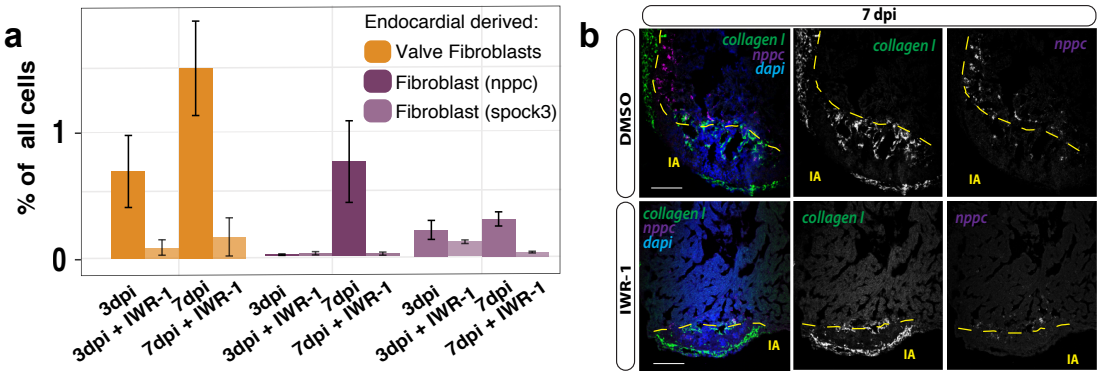


Figure 35. Endocardial derived fibroblasts after Wnt inhibition

a. Changes in abundance of endocardial derived fibroblasts upon Wnt inhibition at 3 and 7 dpi. Error bars show standard error of the mean. **b.** Fluorescent in-situ hybridization of *nppc* fibroblasts (purple) in injured hearts at 7 dpi with or without Wnt inhibition. (scale bar: 100μm)

Depletion of perivascular cells contributes to hypoxia

Another cell type that decreased in number after IWR-1 treatment were the perivascular cells. Based on high expression of the chemokine cxcl12b (see Fig. 23a) and the previously described function of this factor in the context of zebrafish heart regeneration[50], we hypothesized that the perivascular cells could be a major cellular driver of neovascularization, an important event that accompanies regeneration[2,22,50]. In concordance with this hypothesis, we observed that hypoxia markers are up-regulated in Wnt-inhibited hearts upon injury, when perivascular cells were diminished (Fig. 36).

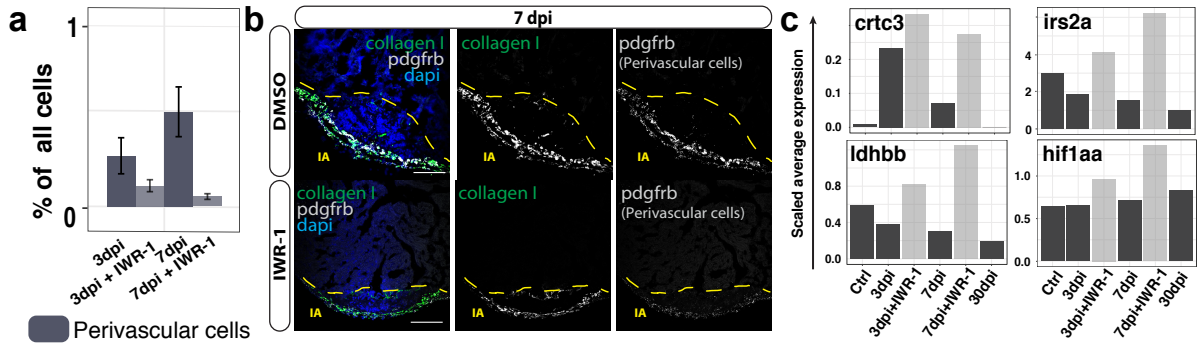


Figure 36. Decrease in perivascular cells led to hypoxia after Wnt inhibition.

a. Changes in abundance of perivascular cells upon Wnt inhibition at 3 and 7 dpi. Error bars show standard error of the mean. **b.** Fluorescent in-situ hybridization of perivascular cells (white) in injured hearts at 7 dpi with or without Wnt inhibition. (scale bar: 100μm) **c.** Expression of hypoxia induced genes at 3 and 7 dpi with and without Wnt inhibition.

Discussion

Simultaneous lineage tracing and transcriptome profiling of single cells

In this project, we developed a new approach for simultaneous lineage tracing and transcriptome profiling of single cells. We achieved this by combining a new experimental protocol repurposing the CRISPR/Cas9 system for lineage barcoding, using scRNA-seq as a readout, and custom tailoring a tree building algorithm to our data. This new approach allowed for the building of developmental lineage trees on top of the single-cell transcriptome data. We could validate our new approach in the zebrafish larvae and in different adult organs. We identified cell types from all three developmental layers from both the developing larva and the adult zebrafish, and our lineage trees successfully reproduced both common (e.g. immune cells from different organs) and split (e.g. endocardium vs. epicardium) lineages of these cell types.

Current droplet-based single-cell sequencing approaches have some inherent imperfections that also affect the performance of our lineage tracing method: Cell duplexes cannot be ruled out due to the stochastic nature of mixing cells and barcoded primer beads using droplet microfluidics. The low capture rate of especially lowly expressed genes can lead to dropouts. Targeted amplification of one selected sequence from the low concentration of single-cell transcriptome cDNA might lead to artificial cell-barcode overwriting due to the exponential nature of the PCR and the high cycle numbers needed. Our new computational algorithm was designed to effectively handle these imperfections, but better experimental procedures including an advance in enzyme chemistry and optimization of the targeted sequence amplification process (e.g. by linear methods instead of PCR) could lead to further improvement of the method.

The cell type atlas of the zebrafish heart in health and disease

With a method that combined scRNA-seq and lineage tracing, we determined the cellular composition of the healthy zebrafish heart and during its regeneration. The unbiased method without cell sorting enabled us to capture all cell types from the zebrafish heart without prior knowledge of their transcriptome and lineage. With the transcriptome information, we could build an atlas of single cells for the zebrafish

heart in health and disease. Additionally, we combined different experimental and in-silico approaches to retrieve spatial and temporal information of the cell types.

In the healthy heart, we found transcriptional differences between cells in the atrium and the ventricle in all three cellular layers of the heart: the epicardium, the myocardium and the endocardium. We confirmed these spatial differences by in-silico mapping of the transcriptome of single cells onto spatially resolved bulk data (Tomo-seq). While spatially differentiated gene expression was described recently for the myocardium[21], it has not been characterized for the other two layers of the heart.

During regeneration, we described a subtype of cardiomyocytes that increased significantly in number at 3 days after injury. Their transcriptome differed from other cardiomyocytes and they expressed genes involved in cardiomyocyte differentiation and development. We suspect these to be undifferentiated or newly formed cardiomyocytes. The change of their location from 3dpi to 7dpi (3dpi: outside injury area, 7dpi: within the injury area) could lead to the hypothesis that they are cardiomyocytes that migrate to the injury site to replace the lost heart muscle cells. Here, further investigation is required.

The role of non-cardiomyocytes in the context of heart regeneration has so far been under-appreciated. In fibroblasts, we observed the expression of known pro-regenerative factors. The sub-clustering of fibroblasts revealed a large diversity of different sub-populations of collagen expressing cell types. These findings challenge the concept of heart fibroblasts being a homogeneous cell type, but rather indicate that the fibrosis and the collagen deposited after the injury have different sources. Most importantly, we found transient cell types that were highly unregulated in number after injury: *col12a1a* expressing fibroblasts, *col11a1a* expressing fibroblasts and *nppc* expressing fibroblasts. *Col12a1a*, a short and non-fibrillar collagen, has been shown to be expressed upon heart injury[53] and was linked to Wnt signaling in the context of zebrafish spinal cord regeneration[54]. *Col11a1a* and *nppc* have not been described in the context of heart regeneration so far. In addition to their transient nature, all of these cell types also produce pro-regenerative factors like *nrg1*, *fn1a*, and retinoic acid, which strongly indicates their pivotal role in regeneration.

Developmental lineages of transient cell types during regeneration

We used the developmental lineage information generated by our new lineage tracing approach to further understand the temporal aspects of these transient fibroblast subtypes. We could establish two lineage origins for fibroblasts: the epicardium and the endocardium. The transient *col11*and *col12* expressing fibroblasts, which we identified as epicardially derived, were shown to transition from constitutive fibroblasts by RNA-velocity. They resemble cell types that were previously described as “activated” fibroblasts in the literature[29]. From our data, they appear to be induced by retinoic acid. From these findings, we conclude that the pro-regenerative function of retinoic acid is at least partially indirect, acting via induction of *col11/col12* fibroblasts, which in turn produce pro-regenerative factors.

While the epicardium was an established developmental origin of cardiac fibroblasts[44], the presence of an endocardially derived fibroblasts population after injury was controversial[29]. We confirmed three populations of endocardium-derived fibroblasts (*nppc* fibroblasts, valve fibroblasts, *spock3* fibroblasts), from which the *nppc* expressing fibroblasts are transient after injury. On top of our lineage data, RNA-velocity also indicated the transition of *nppc* fibroblasts from the ventricular endocardium. We could show that all of these endocardial derived fibroblasts increase in number after injury in a canonical Wnt signaling-dependent manner, which links the endocardial derived fibroblasts to Wnt signaling and strengthens their role in the context of regeneration. Interestingly, the transient *nppc* fibroblasts were only generated upon deep injury that directly affected the endocardium. This could be one reason why they have not been clearly identified so far. We conclude that this is a tissue specific fibrotic response, and probably not indispensable for the regenerative process *per se*, since the heart can also regenerate from a shallow injury without these cells being present.

The role of neovascularization in regeneration

An “activation” of the endocardium (*nppc* fibroblasts) and the increase in number of the perivascular cells (expressing *pdgfrb*) indicated an important role of neovascularization in the context of zebrafish heart regeneration. We found that the perivascular cells are the source of *cxcl12b*, a major factor in neovascularization after heart injury. The source of this factor has not been reported previously[50]. Particularly, when the number of perivascular cells was reduced by Wnt inhibition, an increased hypoxia response could be observed in the regenerating heart. We hypothesize that impaired neovascularization is the main cause of the delayed regeneration upon Wnt inhibition. The role of neovascularization and the cell types involved in this process should be elucidated in further studies.

No evidence of transdifferentiation during zebrafish heart regeneration

In our study, we did not find any evidence of transdifferentiation between e.g. fibroblasts, immune cells and cardiomyocytes. However, we discovered cells that contained genes from different, unrelated cell types (e.g. genes of both macrophages and fibroblasts, fibroblasts and cardiomyocytes etc.). Right now, we think they are likely cell duplexes that were shown to exist in droplet based scRNA-seq approaches[15], because we also observed such phenomena between cell types where a transdifferentiation is very likely impossible (e.g. erythrocytes and fibroblasts). Nevertheless, our data did support that different cell types can be “activated” (*col11/col12* fibroblasts), proliferate, and/or change their transcriptome partly or completely (*nppc* fibroblasts). It can be discussed whether these events create a new cell type or rather a new cell state. So far, there is no clear definition that distinguishes these two terms. We conclude that a definition should consider factors like the degree and length of transcriptional change, lineage similarities, and reversibility of the change.

Conclusion and outlook

Our data provides a detailed, cell type resolved picture of the zebrafish heart both at a healthy state and during the regeneration process. For many known regulatory factors involved in heart regeneration, we were able to show - in many cases for the first time - their corresponding cell types. We also identified crucial cell types that increase in number in the regenerative niche after injury. By developing a new lineage tracing approach, we could capture the developmental lineages of these transient cells. We showed that they are derived from both the epi- and the endocardium. Our experiments further indicated that the transient epicardial fibroblasts are induced by retinoid acid signaling, and the transient endocardial fibroblasts by Wnt signaling, which links two broad acting signaling factors to specific cell types.

To be able to dissect the exact impact of each cell type, further targeted experiments, e.g. by targeted ablation of certain fibroblast subtypes, would be needed. While we focused on the fibrotic response and fibroblast-like cells in heart regeneration, our data also captured different types of immune cells as well as cardiomyocytes, and many of them change in numbers after the injury and during regeneration. Our data can serve as a resource for further studies of these cell types in the future.

The new lineage tracing approach we developed here will be very impactful in the field of single-cell genomics and broadly applicable for various applications. It can be used to investigate the embryonic origin of anatomical structures where the developmental lineage is unclear. In the field of stem cell research, our technology could be used to quantify the different cell types that can be derived from a single stem cell. In more clinically oriented research, our approach can be used to perform lineage tracing of tumors and their metastases, allowing insights into cancer clonality and cancer stem cells. It is possible, however, that the time point of barcode induction has to be optimized to for other applications. Right now, the barcoding events are restricted to the very early stages of the zebrafish development, which limits us to the study of developmental lineages of cell types. For a better control of the barcoding kinetics, chemically inducible Cas9 system[55] or Cas9 under different promotors could be explored. As demonstrated in this work, the CRISPR-based lineage tracing could also be combined with short term lineage determination techniques like RNA-velocity or RNA metabolic labeling[56]. This combination will be useful to resolve both developmental lineages and short-term transitions between cell types.

Methods and materials

Some of these methods were also published in the paper Spanjaard, Hu et al. (2018), Nature Biotechnology[39]. They are included here for completeness.

Zebrafish lines and animal husbandry

The Zebrafish lines used were: wild-type strain *AB,* transgenic lines *Tg[ubi:zebrabow-M]*^{a131Tg} [32], *TgBAC[cryaa:EGFP, tcf21:Cre*^{ERt2}*]*^{pd4} [44], *and Tg[-3.5ubi-loxP-EGFP-loxP-mCherry] (ubi:Switch)* [45]. We did not use the ability of the *zebrabow M line* to switch the fluorescent colors after Cre induction. Adult zebrafish of random sex, between 4 months and one year were used. Zebrafish were bred, raised, maintained, and handled in accordance with the guidelines of the Max Delbrück Center for Molecular Medicine and the Landesamt für Gesundheit und Soziales, Berlin, Germany, for the use of laboratory animals based on the current version of German law on the protection of Animals and EU directive 2010/63/EU on the protection of animals used for scientific purposes. In addition, housing and breeding standards also followed the international “Principles of Laboratory Animal Care” (NIH publication no. 86-23, revised 1985). All animal procedures were approved by the local authorities (LAGeSo, Berlin, Germany) under license number G0211/16.

Injection of one-cell stage embryos

Cas9 protein was purchased from New England Biolabs (NEB). The sgRNA targeting RFP (sequence: GGTGTCCACG TAGTAGTAGC GTTTTAGAGC TAGAAATAGC AAGTTAAAAT AAGGCTAGTC CGTTATCAAC TTGAAAAAGT GGCACCGAGT CGGTGCTTTT) was transcribed *in vitro* using the MEGAscript T7 Transcription Kit (Thermo Scientific). The sgRNA template for in vitro transcription was synthesized with T4 DNA polymerase (NEB) by partially annealing two single-stranded DNA oligonucleotides containing the T7 promoter followed by the RFP binding sequence and the tracrRNA sequence, respectively. For lineage tracing, we set up in-crosses between individuals of the *zebrabow M* line with high RFP fluorescence. We injected the embryos at the one-cell stage with 2 nl Cas9 protein (350 ng/μl) with the sgRNA (50 ng/μl). After 24 hours, we selected embryos with low RFP fluorescence and raised them either to 5dpf or to adulthood until the time point of experiments.

Cryo-injury procedure

Cryo-injury was performed as previously described[3,24]. First, fish were pre-sedated in water containing 0.03 mg/ml Tricaine (PHARMAQ, pH 7). Concentration was then increased to 0.168 mg/ml for anesthesia. Fish were placed with the ventral side facing up into a foam holder under a dissecting scope. To access the heart, a small incision was made through the body wall and the pericardium, using microdissection forceps and scissors. Once the pericardial sac was opened, the heart ventricle was exposed by gently compressing the abdomen. Excess water was carefully removed by blotting with tissue paper, not allowing fish skin to dry. Then, a stainless steel cryoprobe precooled in liquid nitrogen was applied to the ventricular wall. The contact length of the cryoprobe and the heart tissue was 20s, unless mentioned otherwise (for example to access the effect of a shallow vs. deep injury). Fish were then placed in a tank of fresh system water with 1.5mg/l morphine sulphate for 6h. Reanimation was enhanced by the gill oxygenation where water around gills was aerated by pipetting for a couple of minutes.

Preparation of single-cell suspensions

Larvae at 5 dpf: Single larvae were transferred into 50μl HBSS supplemented with 1× TrypLE (Thermo Fisher Scientific) and incubated at 33 °C for ~20 min with intermittent pipette mixing (every 5-7 min) until the larva was no longer visible. To stop the dissociation, 500μl cold HBSS (Thermo Fisher Scientific) containing 1% BSA was added, and the suspension was centrifuged at 4 °C and 300g for 5 min to pellet the cells. The pellet was washed with 500μl cold HBSS containing 0.05% BSA and centrifuged again. The resulting pellet was resuspended in 50μl of the same buffer and filtered through a cell strainer with a mesh of 35-μm diameter.

Adult zebrafish heart: Adult zebrafish were euthanized by an overdose of tricaine in combination with ice water. For the experiments, the hearts were harvested either as a healthy control or at 3, 7, 15, or 30 days post injury. The dissection included the atrium, ventricle and bulbus arteriosus, except for samples in which only the atrium or the ventricle was isolated. A needle and a syringe filled with cold HBSS were used to pierce into the lumen of the heart and thoroughly wash away most of the erythrocytes in the tissue. Afterwards, the tissue was opened carefully with forceps, and the heart tissue was incubated at 37 °C for 30 min in 500μl HBSS containing Liberase enzyme mix (Sigma-Aldrich, 0.26 U/mL final concentration) and Pluronic F-68 (Thermo Fisher

Scientific, 0.1%), while shaking at 750 r.p.m. with intermittent pipette mixing. After most of the tissue was dissociated, the reaction was stopped by adding 500*µ*l cold HBSS supplemented with 1% BSA. The suspension was centrifuged at 250*g* at 4°C and washed two times with 500*µ*l cold HBSS containing 0.05% BSA, then filtered through a cell strainer of 35*µ*m diameter. The quality of the single-cell suspension was then confirmed under the microscope, and cells were counted prior to scRNA-seq library preparation.

Adult zebrafish Brain. The telencephalon without olfactory bulbs was isolated in cold HBSS and immediately transferred to a solution of HBSS with 0.81% d-glucose and 15 mM HEPES. Dissociation was initiated by adding 0.1× TrypLE and 0.1% Pluronic F-68 (final concentrations). The tissue was incubated for 30 min at 37 °C while shaking at 750 r.p.m., with occasional gentle mixing. The dissociation reaction was stopped by the addition of an equal volume of EBSS solution containing 4% BSA and 20 mM HEPES. The suspensions were pelleted, washed and filtered following the procedure described above for the zebrafish heart.

Adult zebrafish Pancreas and liver. The pancreatic tissue containing preferentially the primary pancreatic islet was isolated under a stereomicroscope and transferred into 500*µ*l HBSS containing 1× TrypLE and 0.1% Pluronic F-68. The liver was isolated and dissected into small pieces, one of which was transferred into 500 *µ*l HBSS containing 1× TrypLE and 0.1% Pluronic F-68. After 30 min of incubation at 37 °C with intermittent pipetting, the suspensions were pelleted, washed and filtered following the procedure described above for the zebrafish heart.

Lineage barcode detection in bulk samples

Lineage barcode detection from DNA: DNA of single larvae was extracted by heating the samples in 50*µ*l of 50 mM NaOH at 95 °C for 20 min. 5*µ*l of 1 M Tris-HCl, pH = 8.4 was then added. 20*µ*l of the DNA was used for amplification of lineage barcode sequences using RFP-specific primers; these primers were barcoded to later identify the samples after pooling. The RFP primer design ensured that the target site of Cas9 was positioned in the middle of the sequencing read. We then pooled the PCR products, performed a clean-up reaction using magnetic beads (AMPure Beads, Beckman Coulter), and added Illumina sequencing adapters in a second PCR reaction.

Lineage barcode detection from RNA: RNA of single or pooled animals was extracted with TRIzol Reagent (Thermo Fisher Scientific) according to the manufacturer's protocol. The resulting RNA pellets were resuspended in 10μl of reverse transcription mix (0.3 μM poly-T primer, 1× first strand buffer (Thermo Fisher Scientific), 10 μM DTT, 1 mM dNTPs, 0.5 μl RNAseOUT (Thermo Fisher, Cat. No. 10777019), 0.5 μl SuperScript II (Thermo Fisher, Cat. No. 18064-014). The reaction was incubated at 42 °C for 2 h for reverse transcription, followed by lineage barcode-specific PCR amplification as described above for DNA-based lineage barcode detection.

Determination of lineage barcode probabilities

The reads of DNA-based bulk lineage barcode detection from 32 single embryos were aligned to a reference sequence of RFP. Unmapped reads and reads that did not start with the exact PCR primer were filtered out and all reads were truncated to 100 nucleotides, removing shorter ones. All wildtype RFP reads were removed from each embryo. We normalized the barcode content of each embryo to one and calculated barcode probabilities as the average ratio with which each lineage barcode was observed.

We assigned the bulk probabilities to single-cell lineage barcodes accordingly, when they had the same sequence. To account for the slightly different sequencing read structure of single-cell and bulk lineage detection, we compared only those parts of sequencing reads that are shared between the two approaches (approx. 75 reads in the middle of each sequencing read). Lineage barcodes in single-cell data that were not detected in bulk had their probability set to the lowest probability value detected in bulk.

Determination dynamics of barcode creation

Embryos were injected with Cas9 and sgRNA at the one-cell stage. After 1, 2, 3, 4, 6, 8, 10, and 24 h, several embryos were collected and pooled (5–6 for earlier stages, 2–3 for later stages), followed by RNA and/or DNA extraction using TRIzol Reagent according to manufacturer's protocols. Bulk lineage barcode libraries were produced as described above. For each sample, we calculated the percentage of wildtype sequence RFP. We fit a negative exponential to these data, assuming that the fraction of wildtype sequence RFP at t = 0 was 1.

ScRNA-seq and lineage barcode detection

Single cells were captured using the Chromium Single Cell 3' kit (10X Genomics, PN-1000075), according to the manufacturer's recommendations. During this process, reverse transcription was performed with polyT primers containing cell-specific barcodes, Unique Molecular Identifiers[31] (UMI), and adaptor sequences. We aimed for 10,000 cells per library whenever possible. Both v2- and v3-chemistry were used for data presented here. For simultaneous detection of transcriptome and CRISPR-induced lineage barcodes, the lineage barcode sequences were enriched by a two-round nested PCR approach, using 10μl of the 10X Genomics cDNA. For the nested PCR, two target site specific primers and one reverse primer binding to the adaptor site were used to make sure the barcode and UMI sequence were kept in the lineage reads. Primer sequences are provided under sequences section. The lineage barcode library was then indexed using the indexing primers of the Chromium kit to make sure they could be sequenced and demultiplexed together with the transcriptome. Samples were sequenced on Illumina NextSeq 500 2x 75 bp and Illumina HiSeq 2500 2x 100 bp after successful quality control by Bioanalyzer (DNA HS kit, Agilent).

Mapping and clustering of single-cell mRNA data

A zebrafish transcriptome was created with Cell Ranger 3.0.2 from GRCz11, release 92. Alignment and transcript counting of libraries was done using Cell Ranger 3.0.2. Library statistics are summarized in Additional table 1. The transcriptome data was filtered, clustered, and visualized using Seurat 3.0[8]. and standard parameters. The script and exact parameters for filtering, clustering, and principle component analysis is published in Spanjaard et al.[39]

Mapping and filtering of single-cell lineage data

Lineage barcode filtering was performed as described in Spanjaard et al.[39]. Briefly, lineage barcode sequences were aligned using *bwa mem* to a reference of dTomato. Valid cell barcodes were identified based on the single-cell transcriptome data. We removed reads that were unmapped, had an incorrect cell barcode, or did not start with the exact PCR primer we used. We truncated all lineage barcode sequences to 75 nucleotides and removed shorter sequences. In further filtering steps, we consecutively filtered out lineage barcodes in order to remove all lineage barcodes that did not fulfill the following quality control conditions. As a first step, we required all

molecules to be sequenced at least twice to remove straightforward sequencing errors; this also reduces complexity for the subsequent filtering steps. To remove molecules caused by incorrect annealing, we selected only the most prevalent lineage barcode for each combination of cell barcode and UMI, only the most prevalent UMI for each cell barcode and lineage barcode, and only the most prevalent cell barcode for each UMI and lineage barcode. We then assessed similar (Hamming distance of 2 or less) lineage barcode sequences in each cell, only keeping both if they obeyed criteria designed to test whether both sequences are indeed correct (see Spanjaard et al. for details). Finally, we removed cells that looked like doublets by comparing their number of different lineage barcodes to the lineage barcode number distribution for that cell type.

Histological staining, analysis and imaging

For the analysis of the injury areas, animals were humanely killed at different times post-injury by placing them in ice cold water of 0-4°C for 20 minutes. Hearts were dissected and incubated in 2U/ml heparin and 0.1KCl in PBS for ca 30min. Cryosamples were fixed in 4% PFA in PBS overnight at 4°C, washed in PBS for 3 x 10 min and incubated overnight in 30% sucrose in PBS. Samples were then frozen in Tissue-Tek O.C.T Compound (Sakura) on dry ice. Tissue was cut at 7µm on a cryostat (Leica) using Superfrost slides (ThermoFisher Scientific). Connective tissue was stained using acid fuchsin orange G (AFOG). In brief, slides were dried for 30min at RT. Next, slides were incubated in Bouin Solution (Sigma-Aldrich) for 2h at 60°C and left for overnight incubation under the hood. Slides were washed for 30min under running ddH_2O and incubated for 7min in 1% phosphomolybdic acid (Sigma-Aldrich). Samples were washed for 3min in running ddH_2O and incubated with AFOG solution (self-made, Sigma-Aldrich) for 3min. Slides were washed until clear with running ddH_2O and rehydrated with 70%, 94%, 2x 100% ethanol and 2x 5min with xylol (Sigma-Aldrich). Slides were mounted with xylene mounting medium (Merck) and let dry overnight under the hood. For the analysis of injury size, the total ventricular tissue area and injury area (IA) on multiple slides were measured. Imaging was done using a Keyence Microscope BZX800 and analyzed with ImageJ/Fiji.

Immunofluorescence *in situ* hybridization and imaging

For fluorescent *in situ* hybridization, we used the RNAscope assay (ACD) following the manufacturer's instructions for fixed frozen tissues (323100-USM). The 7μm cryo-sections were dried for 30min at RT before the experiment. The following changes were made: slides were incubated at 99°C for 15min with an antigen retrieval solution (ACD) using a steamer (WMH). Signal development was done using TSA plus fluorophores (PerkinElmer) fluorescein, Cyanine3 and Cyanine5 in 1:1000 dilution. Probes used are listed in the Additional table 6. Imaging was done using Zeiss LSM880 confocal microscopy and analysis was performed using ImageJ/Fiji and Photoshop software.

Wnt inhibition

For Wnt inhibition experiments, the Wnt antagonist IWR-1-endo (Sigma-Aldrich) was used[54]. IWR-1 was dissolved in DMSO to prepare a stock concentration of 10mM. Wild-type fish were injected intraperitoneally with 25µl of 10µM IWR-1 in PBS or DMSO (0.1% in PBS)[57]. The injection was performed at 1dpi, 2dpi for the 3dpi analysis and for later time points, once every two days from 2dpi until the day of sacrifice.

Cre/lox lineage tracing

For Cre/lox lineage tracing *TgBAC[cryaa:EGFP, tcf21:Cre^{ERt2}]^{pd42}; Tg[-3.5ubi-loxP-EGFP-loxP-mCherry]* transgenic fish were treated with 10µM 4-OHT (Sigma-Aldrich) dissolved in 100ml of water per fish, 4 and 3 days before the cryo-injury for 12h as described[45]. Before the systemic administration, the 10 mM stock (dissolved in ethanol) was heated for 10min at 65 °C. Hearts were harvested 7 dpi.

Tree building

Tree building is done as described in Spanjaard et al.[39]. In short, this lineage tree building algorithm is designed to build lineage trees of high numbers of cells, despite the low detection rate inherent in single-cell sequencing. It consists of two steps: we first determine the lineage tree in terms of lineage barcodes created, and afterwards place the single cells in their appropriate positions in the tree.

To build the barcode lineage we use the dataset of lineage barcodes in single cells and determine which lineage barcodes can be found in the same branch of the lineage trees: those lineage barcodes are observed together in the same cell. We create a graph of lineage barcodes with connections between lineage barcodes indicating that they occur together in a branch. A lineage barcode that has been created early is present in many branches and is connected to many other lineage barcodes. As a corollary, the lineage barcode that is connected to most other lineage barcodes in this graph has been created before all other lineage barcodes. This lineage barcode becomes the root of the lineage tree; it is removed from the lineage barcode graph and we identify the most highly-connected lineage barcode in the resulting graph. This lineage barcode is placed underneath the first lineage barcode in the lineage tree and is again removed from the graph. If at any point the graph splits into disconnected parts after removing a lineage barcode, those parts are distinct branches of the lineage tree and are treated separately. This process is repeated until all lineage barcodes have been placed in the lineage tree.

The above tree construction algorithm is valid for high, non-variable, lineage barcode detection rates, perfect filtering and no doublets; for more realistic scenarios, computations are more complicated but follow the general scheme as outlined above, they are referred to in Spanjaard et al.[39] in detail.

Once the shape of the lineage tree and the location of all lineage barcodes in the lineage tree have been determined, we place the cells in the tree. Cells are placed using the lineage barcodes that are found in them, at the lowest possible position in the tree. Importantly, this means that some cells may be placed in leaf nodes, but typically some cells in which a leaf node lineage barcode has not been detected will be placed in internal nodes.

Lineage determination

To determine lineage relationships between cell types, we calculated cell type ratios - the amount of cells of a specific type divided by the node size - for each node, and calculated the lineage similarity between two cell types as the correlation between their cell type ratios, weighted by the size of the nodes. Importantly, this approach allowed us to integrate lineage data for multiple biological replicates. Using cell type ratios emphasizes nodes that are lower in the trees and contain fewer cell types, and weighting by node size emphasizes larger nodes to maximize reproducibility.

We determined the stability of clusters by a fifty-fold repetition of downsampling all lineage trees to half of the cells, calculating the correlations and clustering the cell types. At four clusters for 3dpi and seven clusters for 7dpi, we recovered the same clusters in the majority of cases, indicating that these clusters cannot be attributed to sampling artifacts (Additional Figure 10c and 10d).

To determine possible sources of *nppc* fibroblasts, *spock3* fibroblasts and valve fibroblasts, we determined the overlap of cell types with these target cell types in tree leaf nodes. The source-target ratio was calculated as the number of leaf nodes that contained cells of both source and target type, divided by the number of leaf nodes that contained cells of the target type. For example, at 7dpi, ventricular endocardium has a source-target ratio of 1 with *nppc* fibroblasts, meaning that at 7dpi, every leaf node that contains *nppc* fibroblasts also contains ventricular endocardium.

Trajectory analysis and RNA velocity

We used scvelo[46] for integrated trajectory and RNA velocity analysis. First, we extracted velocity information from reads aligned to a transcriptome based on GRCz11, release 92, using Cell Ranger 3.1 and velocyto[46]. Using cell type annotations determined previously, we selected relevant cells, regressed out total counts, number of genes, and mitochondrial percentages, determined highly variable genes and principal components, and used bbknn to integrate datasets. We then reclustered the data using the leiden algorithm. For the integrated dataset, we calculated RNA velocities using the stochastic model implemented in scvelo and plotted those on the UMAP representation of the data.

Tomo-seq Data Deconvolution

We used AutoGeneS[42] (v1.0, https://github.com/theislab/AutoGeneS) to deconvolve the Tomo-seq data. As input, the single-cell data were subset to the 5,000 highly variable genes using SCANPY[58] (pp.highly_variable_genes). AutoGeneS then selected in total 400 informative genes from the highly variable ones that differentiated the cell types. The highly variable genes were selected based on normalized dispersion. The proportions were inferred using non-negative least squares[59] (scipy.optimize.nnls) using the cellular mean expression of the informative genes. For each bulk sample, negative proportions were set to zero and the rest were normalized to sum to one.

Materials

Chemical or reagent	Composition	Order number	Company
Affi Gel blue media polyacrylamide beads	-	-	BIORAD
Agarose	-	840004	Biozym
AMPure XP beads	-	A63880	Beckmann Coulter
BSA	-	A9418	Sigma-Aldrich
Chloroform	-	7331.2	Carl Roth
DBPS	-	-	Gibco
DNA Pol I, E. coli	-	M0209L	NEB
dNTPs	-	R0181	Thermo
DTT (cell suspensions)	-	6908.2	Carl Roth
DTT 0.1 M	-	-	Invitrogen
EDTA 0.5 M solution	-	90004931	Molecula
ERCC spike-in RNA	-	4456740	Ambion
Ethanol	-	9065.2	Carl Roth
Exo I	-	M0293L	NEB
First strand buffer (part of the Kit)	-	-	Invitrogen
Fragmentation buffer	200 mM Tris-acetate, pH 8.1, 0.5 M KOAc, 0.15 M MgOAc		
Gene Ruler DNA ladder	-	SM1331	Thermo
Glycoblue	-	M9516	Ambion
HBSS	-	14175095	Life technologies
Isopropanol	-	9866.2	Carl Roth
Liberase	-	5401119001	Sigma-Aldrich
Methanol	-	11646935	Th. Geyer
NEBnext High Fidelty 2x PCR Master mix	-	M0541L	NEB
NEBnext ultra II End repair/ A-tailing module	-	E7546S	NEB
NEBnext ultra II ligation module	-	E7595	NEB
RNA clean XP beads	-	A63987	Beckman Coulter
RNaseOUT	-	10777019	Life Technologies
rSAP	-	M0371L	NEB
Second strand buffer	-	10812014	Life Technologies
SP6 polymerase	-	M0207	NEB
Superscript II	-	18064014	Invitrogen
Superscript III	-	18080085	Invitrogen
TBE buffer	-	42557.01	Serva
TRIzol reagent	-	AM9516	Ambion
HiScribe T7 ARCA mRNA Kit (with tailing)	-	E2060	NEB
Chromium Single Cell 3' Reagent Kits	-	PN-1000075	10X Genomics
NextSeq 500/550 HO Kit v2.5 (150 Cycles)	-	20024907	Illumina

Consumables		**Order number**	**Company**
Toothpicks		-	Local supermarket
DNA LoBind tubes 0.5 mL		30108035	Eppendorf
DNA LoBind tubes 1.5 mL		30108078	Eppendorf
Falcon tubes, 15 mL		TPP91015	Faust
Falcon tubes, 50 mL		TPP91051	Faust
PCR stripes		A1402-3700	Starlab
Equipment		**Model**	**Company**
Cryotome		Cryo Star NX70	Thermo Scientific
Thermocycler		Nexus gradient/nexus X2e	Eppendorf
Table top centrifuges		5427R	Eppendorf
Heat block		ThermoStat C	Eppendorf
Magnetic separator		For PCR Tube and 1.5/2.0 Tube	10x
Qubit fluorometer		Qubit 4	Invitrogen
Bioanalyzer		2100 Bioanalyzer	Agilent
Vortex		Vortex-Genie 2	Carl Roth
Gel chamber		EasyPhor Mini/Medi	Biozym
Pipette set		Research Plus Serie	Eppendorf
Aluminium rack		Alu Rack 1.5/2/5	Carl Roth

Appendix

Bibliography

1. K. D. Poss, Heart Regeneration in Zebrafish. *Science*. **298**, 2188–2190 (2002).
2. J. M. Gonzalez-Rosa, C. E. Burns, C. G. Burns, Zebrafish heart regeneration: 15 years of discoveries. *Regeneration*. **4**, 105–123 (2017).
3. González-Rosa, J. M., & Mercader, N. (2012). Cryoinjury as a myocardial infarction model for the study of cardiac regeneration in the zebrafish. Nature Protocols, 7(4), 782–788.
4. K. Kikuchi, J. E. Holdway, A. A. Werdich, R. M. Anderson, Y. Fang, G. F. Egnaczyk, T. Evans, C. A. MacRae, D. Y. R. Stainier, K. D. Poss, Primary contribution to zebrafish heart regeneration by gata4+ cardiomyocytes. *Nature*. **464**, 601–605 (2010).
5. C. Jopling, E. Sleep, M. Raya, M. Martí, A. Raya, J. C. I. Belmonte, Zebrafish heart regeneration occurs by cardiomyocyte dedifferentiation and proliferation. *Nature*. **464**, 606–609 (2010).
6. F. Tang et al., mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods. 6, 377–382 (2009).
7. A. K. Shalek et al., Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature. 498, 236–240 (2013).
8. B. Treutlein et al., Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature, 1–16 (2014).
9. T. Hashimshony, M. Feder, M. Levin, B. K. Hall, I. Yanai, Spatiotemporal transcriptomics reveals the evolutionary history of the endoderm germ layer. Nature, 1–18 (2014).
10. Q. Deng, D. Ramsköld, B. Reinius, R. Sandberg, Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. Science. 343, 193–196 (2014).
11. D. A. Jaitin et al., Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. Science. 343, 776–779 (2014).
12. Rosenberg, A. B. et al. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. Science 360, 176–182 (2018).
13. E. Z. Macosko et al., Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell. 161, 1202–1214 (2015).
14. A. M. Klein et al., Droplet Barcoding for Single-Cell Transcriptomics Applied to Embryonic Stem Cells. Cell. 161, 1187–1201 (2015).
15. Zheng, G. X. Y. et al. Massively parallel digital transcriptional profiling of single cells. Nat Comms 8, 14049 (2017).
16. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol 36, 411–420 (2018).
17. Stuart, T. et al. Comprehensive Integration of Single-Cell Data. Cell 177, 1888–1902.e21 (2019).
18. A. Zeisel et al., Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science. 347, 1138–1142 (2015).

19. D. Grün et al., Single-cell messenger RNA sequencing reveals rare intestinal cell types. Nature. 525, 251–255 (2015).

20. Tucker, N. R. *et al.* Transcriptional and Cellular Diversity of the Human Heart. *Circulation* (2020).

21. Litviňuková, M. *et al.* Cells of the adult human heart. *Nature* (2020)

22. Farbehi, N. *et al.* Single-cell expression profiling reveals dynamic flux of cardiac stromal, vascular and immune cells in health and injury. *eLife* **8**, e43882 (2019).

23. Honkoop, H. *et al.* Single-cell analysis uncovers that metabolic reprogramming by ErbB2 signaling is essential for cardiomyocyte proliferation in the regenerating heart. *eLife* **8**, e50163 (2019)

24. J. M. Gonzalez-Rosa, V. Martin, M. Peralta, M. Torres, N. Mercader, Extensive scar formation and regression during heart regeneration after cryoinjury in zebrafish. *Development*. **138**, 1663–1674 (2011).

25. Y. Fang, V. Gupta, R. Karra, J. E. Holdway, K. Kikuchi, K. D. Poss, Translational profiling of cardiomyocytes identifies an early Jak1/Stat3 injury response required for zebrafish heart regeneration. *Proceedings of the National Academy of Sciences*. **110**, 13416–13421 (2013).

26. J. Münch, D. Grivas, Á. González-Rajal, R. Torregrosa-Carrión, J. L. de la Pompa, Notch signalling restricts inflammation and *serpine1* expression in the dynamic endocardium of the regenerating zebrafish heart. *Development*. **144**, 1425–1440 (2017).

27. A.-S. de Preux Charles, T. Bise, F. Baier, J. Marro, A. Jaźwińska, Distinct effects of inflammation on preconditioning and regeneration of the adult zebrafish heart. *Open Biol.* **6**, 160102 (2016).

28. S.-L. Lai, R. Marín-Juez, P. L. Moura, C. Kuenne, J. K. H. Lai, A. T. Tsedeke, S. Guenther, M. Looso, D. Y. Stainier, Reciprocal analyses in zebrafish and medaka reveal that harnessing the immune response promotes cardiac regeneration. *eLife*. **6**, e25605 (2017).

29. H. Sánchez-Iranzo, M. Galardi-Castilla, A. Sanz-Morejón, J. M. González-Rosa, R. Costa, A. Ernst, J. Sainz de Aja, X. Langa, N. Mercader, Transient fibrosis resolves via fibroblast inactivation in the regenerating zebrafish heart. *Proc Natl Acad Sci USA*. **115**, 4188–4193 (2018).

30. Ozhan, G. & Weidinger, G. Wnt/β-catenin signaling in heart regeneration. *Cell Regeneration* **4**, 4:3 (2015).

31. Moore-Morris, T., Guimarães-Camboa, N., Yutzey, K. E., Pucéat, M., & Evans, S. M. Cardiac fibroblasts: from development to heart failure. Journal of Molecular Medicine. 93(8), 823–830 (2015).

32. Y. A. Pan et al., Zebrabow: multispectral cell labeling for cell tracing and lineage analysis in zebrafish. Development. 140, 2835–2846 (2013).

33. F. Amat et al., Fast, accurate reconstruction of cell lineages from large-scale fluorescence microscopy data. Nat Methods. 11, 951–958 (2014).

34. P.J. Keller, Imaging morphogenesis: technological advances and biological insights. Science. 340, 1234168 (2013).

35. T. N. M. Schumacher, C. Gerlach, J. W. J. van Heijst, Mapping the life histories of T cells. Nat Rev Immunol. 10, 621–631 (2010).

36. R. Lu, N. F. Neff, S. R. Quake, I. L. Weissman, Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding. Nature Biotechnology. 29, 928–933 (2011).

37. S. H. Naik et al., Diverse and heritable lineage imprinting of early haematopoietic progenitors. Nature. 496, 229–232 (2013).

38. Kivioja, T. et al. Counting absolute numbers of molecules using unique molecular identifiers. Nat Meth 9, 72–74 (2011).

39. B. Spanjaard, B. Hu, N. Mitic, P. Olivares-Chauvet, S. Janjuha, N. Ninov, J. P. Junker, Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars. Nature Biotech. 36, 469–473 (2018).

40. McKenna, A. et al. Whole organism lineage tracing by combinatorial and cumulative genome editing. Science aaf7907 (2016).

41. J. P. Junker, E. S. Noël, V. Guryev, K. A. Peterson, G. Shah, J. Huisken, A. P. McMahon, E. Berezikov, J. Bakkers, A. van Oudenaarden, Genome-wide RNA Tomography in the Zebrafish Embryo. Cell. 159, 662–675 (2014).

42. H. Aliee, F. Theis, “AutoGeneS: Automatic gene selection using multi-objective optimization for RNA-seq deconvolution” (preprint, Bioinformatics, 2020).

43. Gemberling, M., Karra, R., Dickson, A. L. & Poss, K. D. Nrg1 is an injury-induced cardiomyocyte mitogen for the endogenous heart regeneration program in zebrafish. eLife 4, e05871 (2015).

44. K. Kikuchi, V. Gupta, J. Wang, J. E. Holdway, A. A. Wills, Y. Fang, K. D. Poss, tcf21+ epicardial cells adopt non-myocardial fates during zebrafish heart development and regeneration. Development. 138, 2895–2902 (2011).

45. C. Mosimann, C. K. Kaufman, P. Li, E. K. Pugach, O. J. Tamplin, L. I. Zon, Ubiquitous transgene expression and Cre-based recombination driven by the ubiquitin promoter in zebrafish. Development. 138, 169–177 (2011).

46. V. Bergen, M. Lange, S. Peidli, F. A. Wolf, F. J. Theis, Generalizing RNA velocity to transient cell states through dynamical modeling. Nat Biotechnol (2020), doi:10.1038/s41587-020-0591-3.

47. G. La Manno, R. Soldatov, A. Zeisel, E. Braun, H. Hochgerner, V. Petukhov, K. Lidschreiber, M. E. Kastriti, P. Lönnerberg, A. Furlan, J. Fan, L. E. Borm, Z. Liu, D. Bruggen, J. Guo, X. He, R. Barker, E. Sundström, G. Castelo-Branco, P. Cramer, I. Adameyko, S. Linnarsson, P. V. Kharchenko, RNA velocity of single cells. Nature. 560, 1–25 (2018).

48. J. K. Chen, J. Taipale, K. E. Young, T. Maiti, P. A. Beachy, Small molecule modulation of Smoothened activity. Proc Natl Acad Sci USA. 99, 14071–14076 (2002).

49. L. Zhao, R. Ben-Yair, C. E. Burns, C. G. Burns, Endocardial Notch Signaling Promotes Cardiomyocyte Proliferation in the Regenerating Zebrafish Heart through Wnt Pathway Antagonism. Cell Reports. 26, 546-554.e5 (2019).

50. R. Marín-Juez, H. El-Sammak, C. S. M. Helker, A. Kamezaki, S. T. Mullapuli, S.-I. Bibli, M. J. Foglia, I. Fleming, K. D. Poss, D. Y. R. Stainier, Coronary Revascularization During Heart Regeneration Is Regulated by Epicardial and Endocardial Cues and Forms a Scaffold for Cardiomyocyte Repopulation. Developmental Cell. 51, 503-515.e4 (2019).

51. K. Kikuchi, J. E. Holdway, R. J. Major, N. Blum, R. D. Dahn, G. Begemann, K. D. Poss, Retinoic Acid Production by Endocardium and Epicardium Is an Injury Response Essential for Zebrafish Heart Regeneration. Developmental Cell. 20, 397–404 (2011).

52. J. Wang, R. Karra, A. L. Dickson, K. D. Poss, Fibronectin is deposited by injury-activated epicardial cells and is necessary for zebrafish heart regeneration. *Developmental Biology*. **382**, 427–435 (2013).

53. J. Marro, C. Pfefferli, A.-S. de Preux Charles, T. Bise, A. Jaźwińska, Collagen XII Contributes to Epicardial and Connective Tissues in the Zebrafish Heart during Ontogenesis and Regeneration. PLoS ONE. 11, e0165497 (2016).

54. D. Wehner, T. M. Tsarouchas, A. Michael, C. Haase, G. Weidinger, M. M. Reimer, T. Becker, C. G. Becker, Wnt signaling controls pro-regenerative Collagen XII in functional spinal cord regeneration in zebrafish. Nat Commun. 8, 126 (2017).

55. Campbell, L. J., Willoughby, J. J., & Jensen, A. M. Two Types of Tet-On Transgenic Lines for Doxycycline-Inducible Gene Expression in Zebrafish Rod Photoreceptors and a Gateway-Based Tet- On Toolkit. Plos One, 7(12), e51270 (2012).

56. V. A. Herzog et al., Thiol-linked alkylation of RNA to assess expression dynamics. Nature Methods. 14, 1198–1204 (2017)

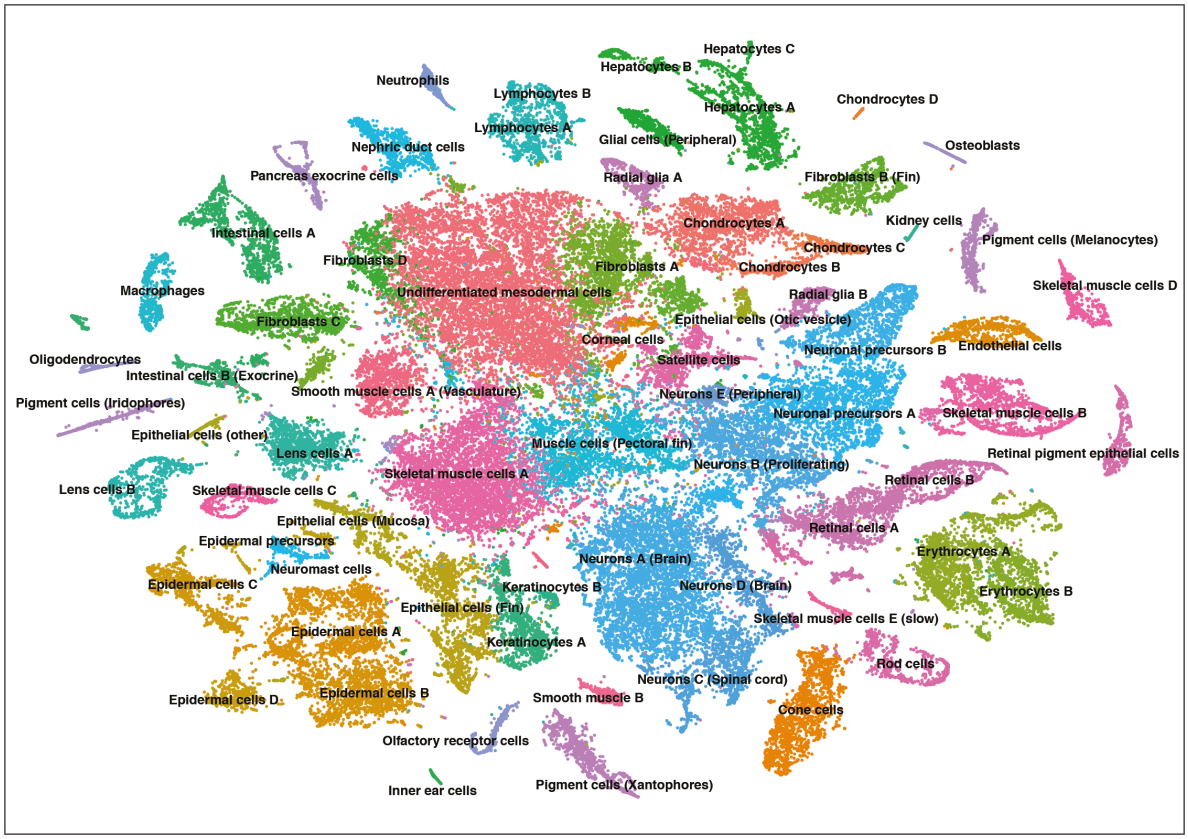
57. M. D. Kinkel, S. C. Eames, L. H. Philipson, V. E. Prince, Intraperitoneal Injection into Adult Zebrafish. JoVE, 2126 (2010).

58. F. A. Wolf, P. Angerer, F. J. Theis, SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. 19, 15 (2018).

59. Lawson C., Hanson R.J., Solving Least Squares Problems, SIAM (1987).

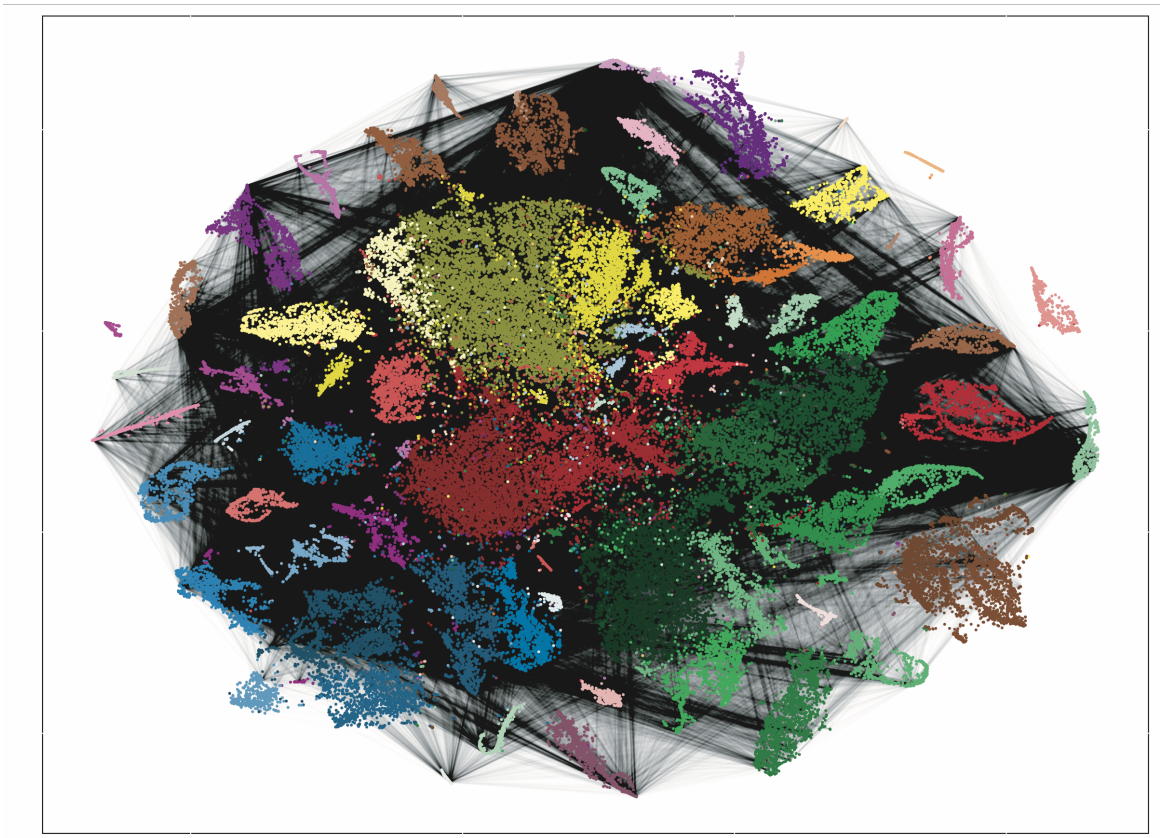
Additional Figures

Additional Figure 1.



Additional Figure 1. Cell type identification in 5 dpf zebrafish larvae (n=7 animals).
tSNE representation of clustering results for scRNA-seq data. Labels indicate cell type assignment.

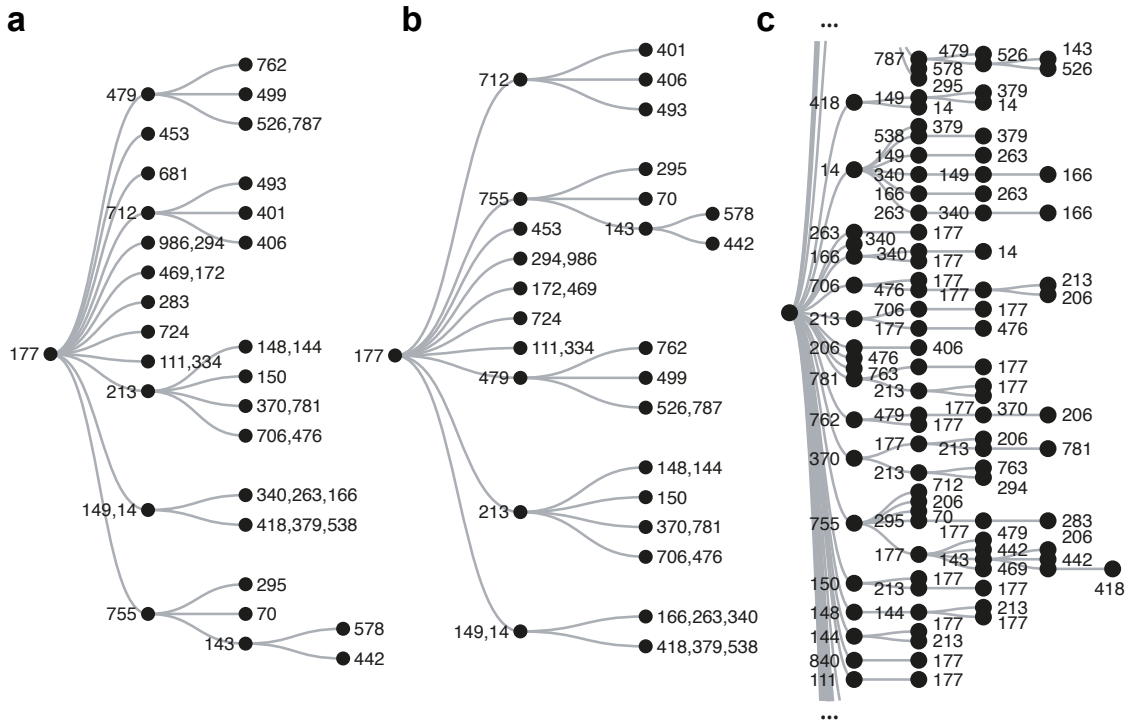
Additional Figure 2.



Additional Figure 2. Connections between pairs of cell types by lineage barcodes.

In this t- SNE representation of scRNA-seq data for dissociated zebrafish larvae (see also Results, Figure 8), pairs of cells with at least one common lineage barcode are linked by black lines. Lineage barcodes with creation probability ≥0.001 and those that were detected in more than 1 larva were excluded from the analysis.

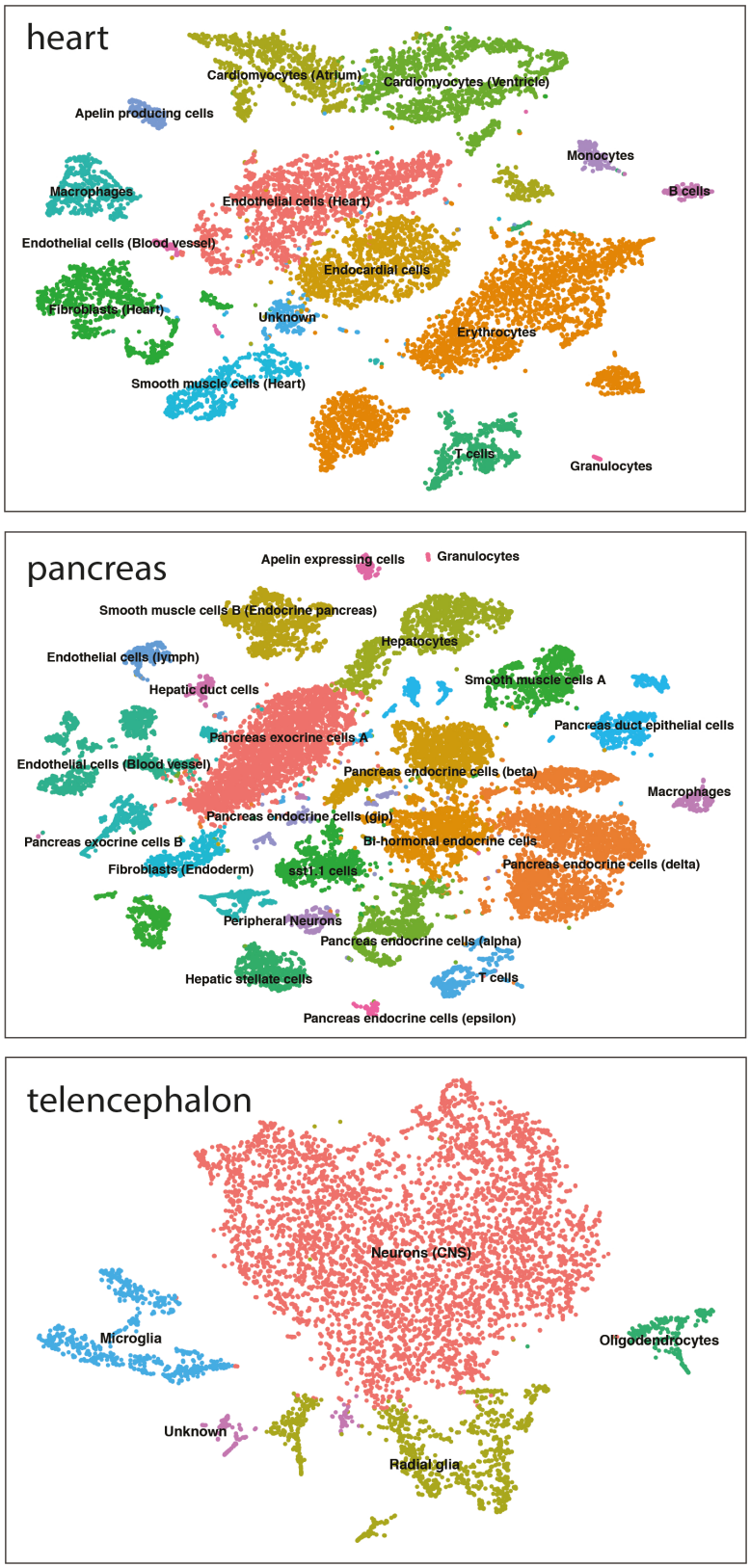
Additional Figure 3.



Additional Figure 3. Comparison of LINNAEUS tree building to maximum parsimony on simple model tree.

a. Simulated lineage barcode tree with realistic parameters (e.g. dropouts and lowly expressed scars). Each number represent a barcode sequence. b. The LINNAEUS tree building algorithm accurately reconstructed the tree. c. Maximum parsimony tree (only a part is shown here). Maximum parsimony did not reconstruct the lineage tree. It introduced considerable amounts of unjustified complexity. Overall, the reconstructed tree contains 46 barcodes and 265 barcode creation events, which means each barcode is placed on average more than 5 times in the tree.

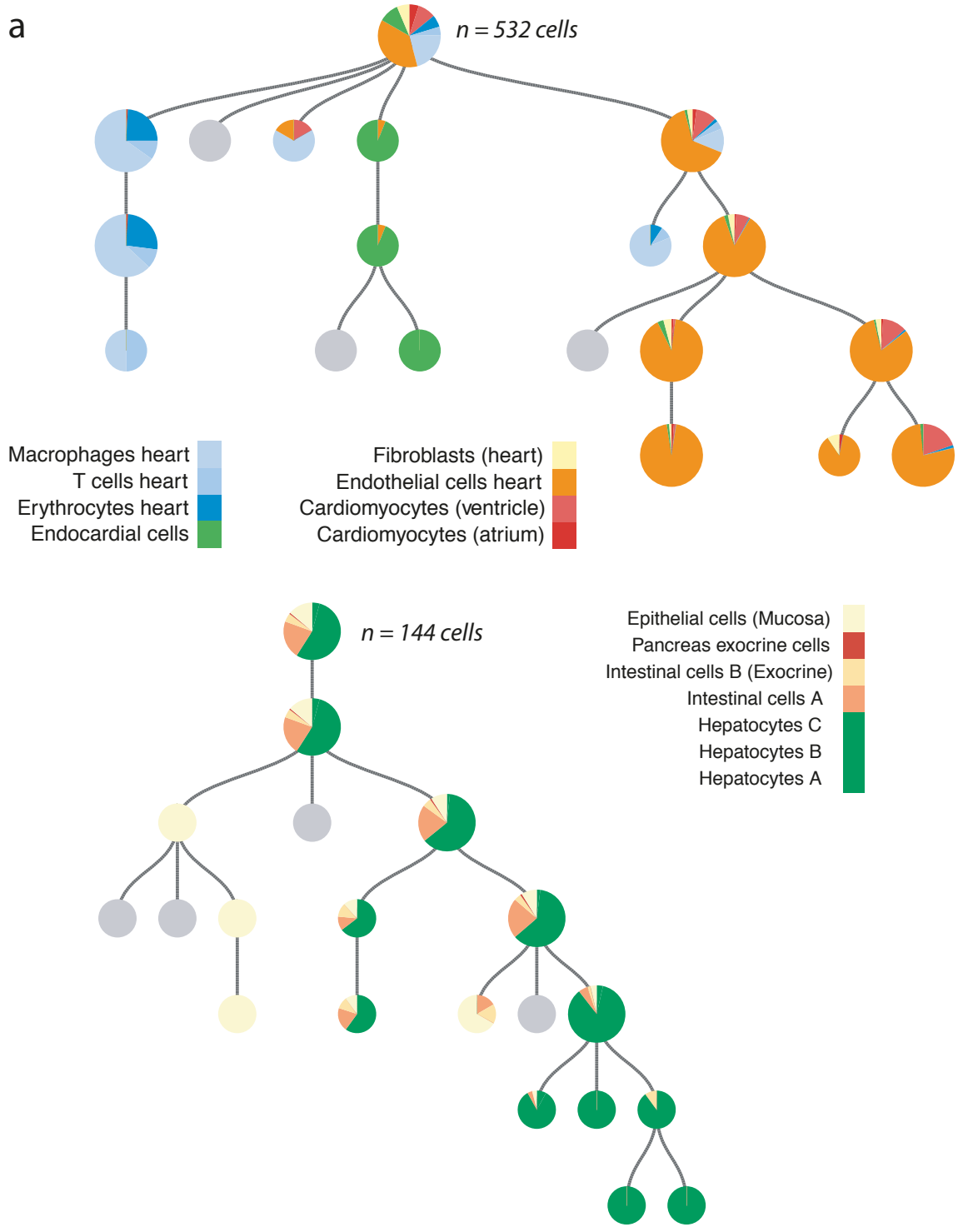
Additional Figure 4.



Additional Figure 4. Identified cell types in adult zebrafish organs.

tSNE representation of clustering results for scRNA-seq data in dissociated heart, pancreas and telecephalon (n=3 animals). Labels indicate cell type assignment.

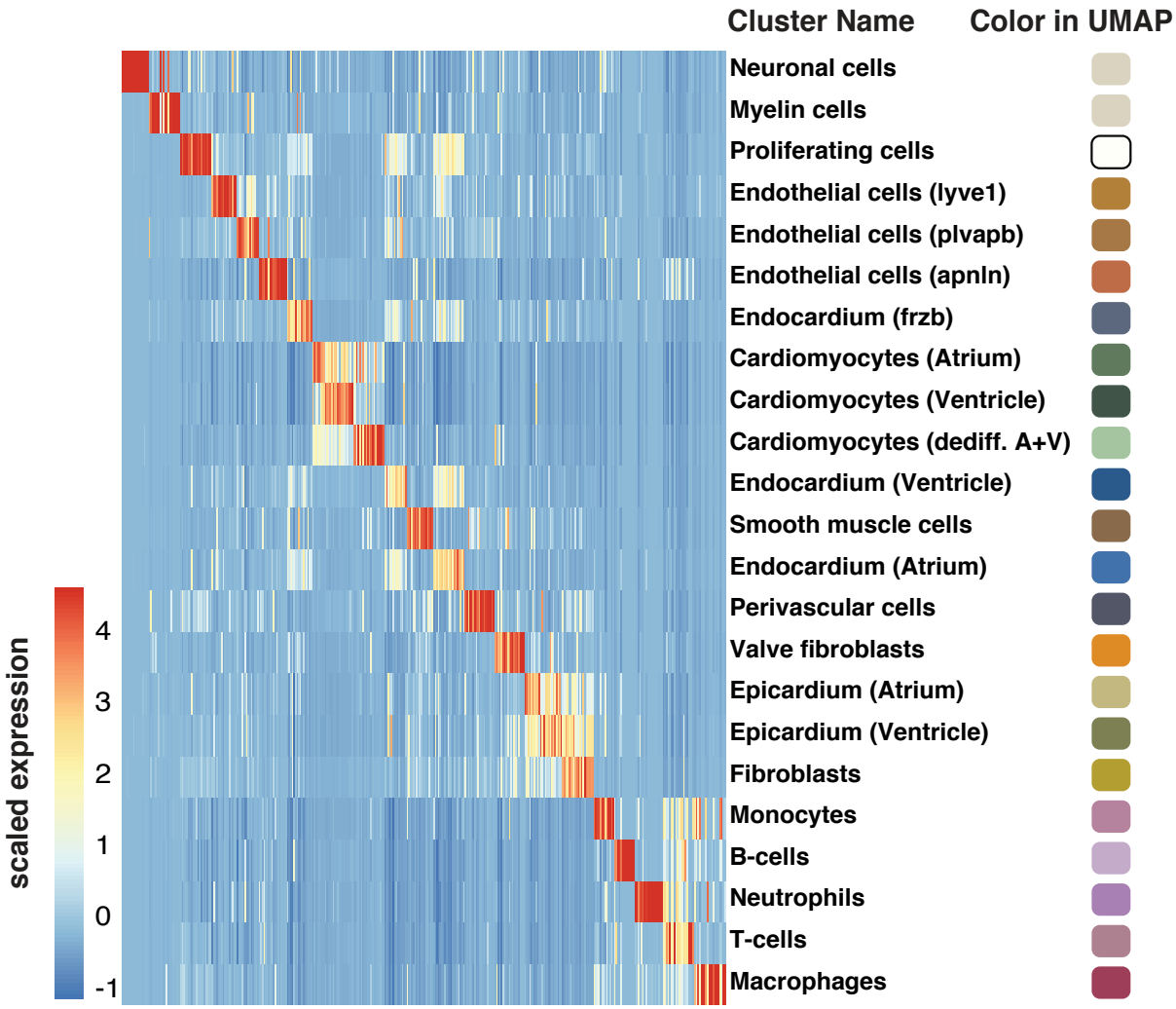
Additional Figure 5.



Additional Figure 5. Zoom into cardiac and pancreatic lineages for one adult.

a. Heart lineage tree. We observed an early separation of myocardial and endocardial lineages. **b.** Liver and pancreas lineage tree. We could observe a split between the intestinal exocrine cells and the hepatocytes.

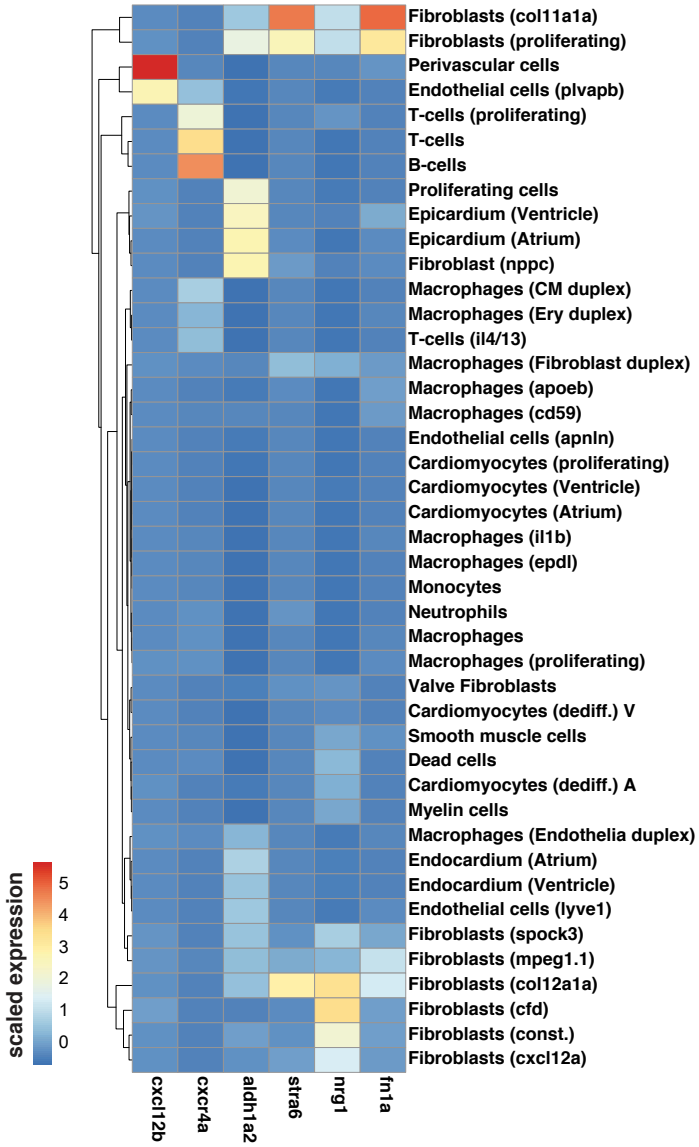
Additional Figure 6.



Additional Figure 6. Top differentially expressed genes for each cell type shown in Fig. 17.

The genes used in this heatmap can be found in Additional table 4. The legend on the right shows the color of the respective cell type in Fig. 17.

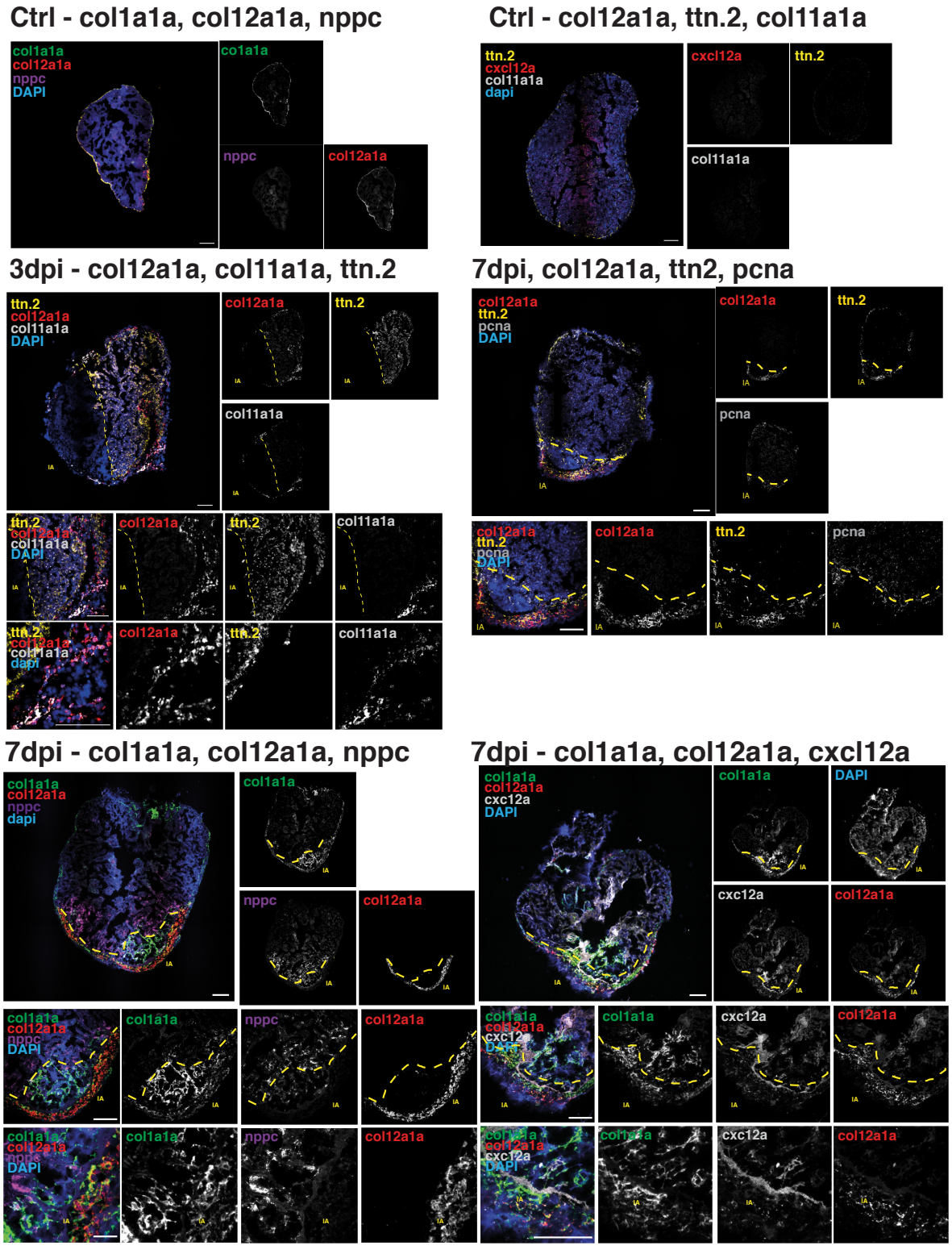
Additional Figure 7.



Additional Figure 7. Expression of known pro-regenerative factors across all cell types in the zebrafish heart.

Expression patterns of cxcl12b, cxcr4a, aldh1a2, stra6, nrg1, and fn1a are show for all cell types and sub-cell types found in the dataset.

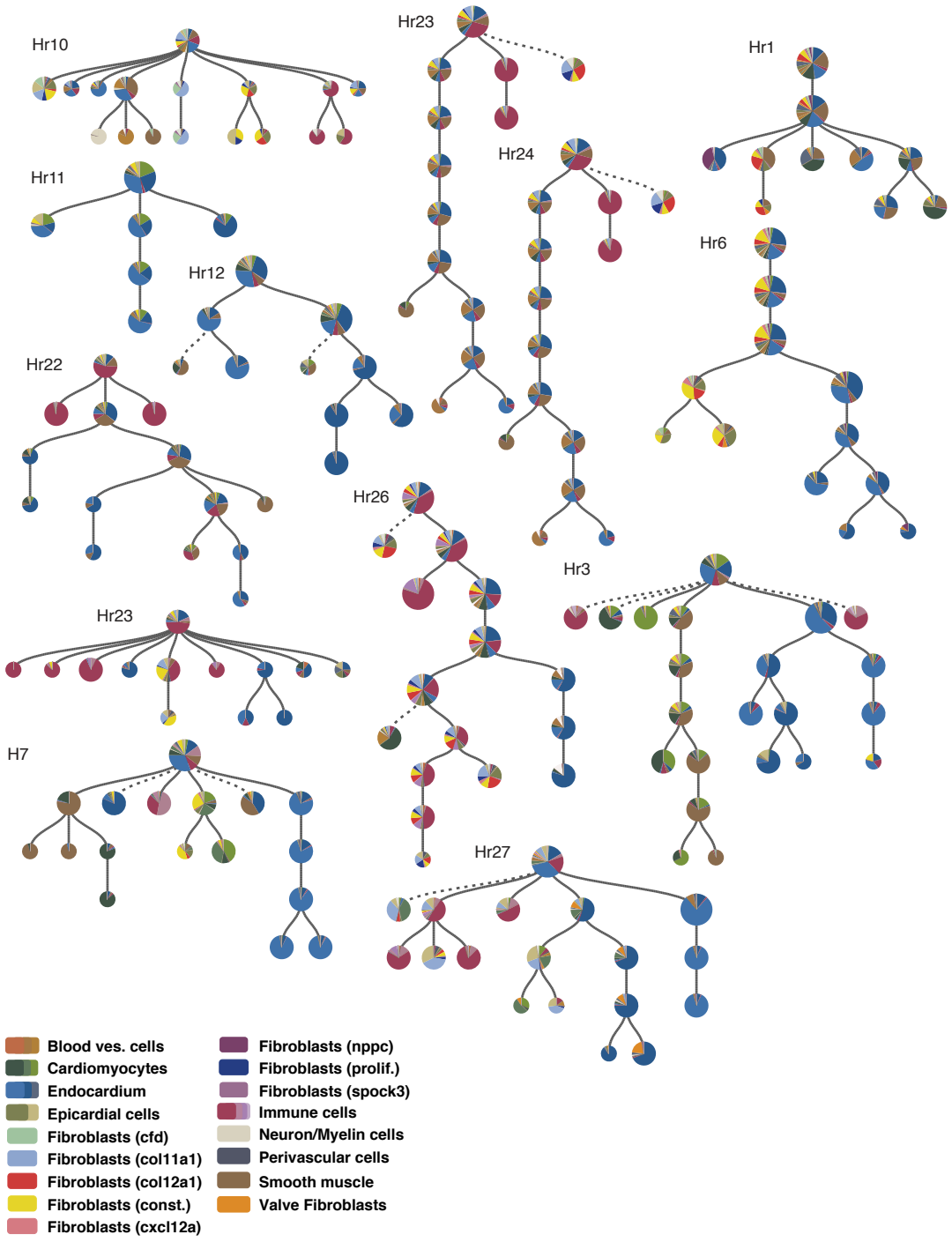
Additional Figure 8.



Additional Figure 8. Fluorescent in-situ hybridization of marker genes for cell types of interest at different time points of injury.

Markers for transient cell types (*col12a1a*, *col11a1a*, *ttn.2*, *nppc*) are localized at the injury area (IA, dotted yellow line). Scale bar: 100μm.

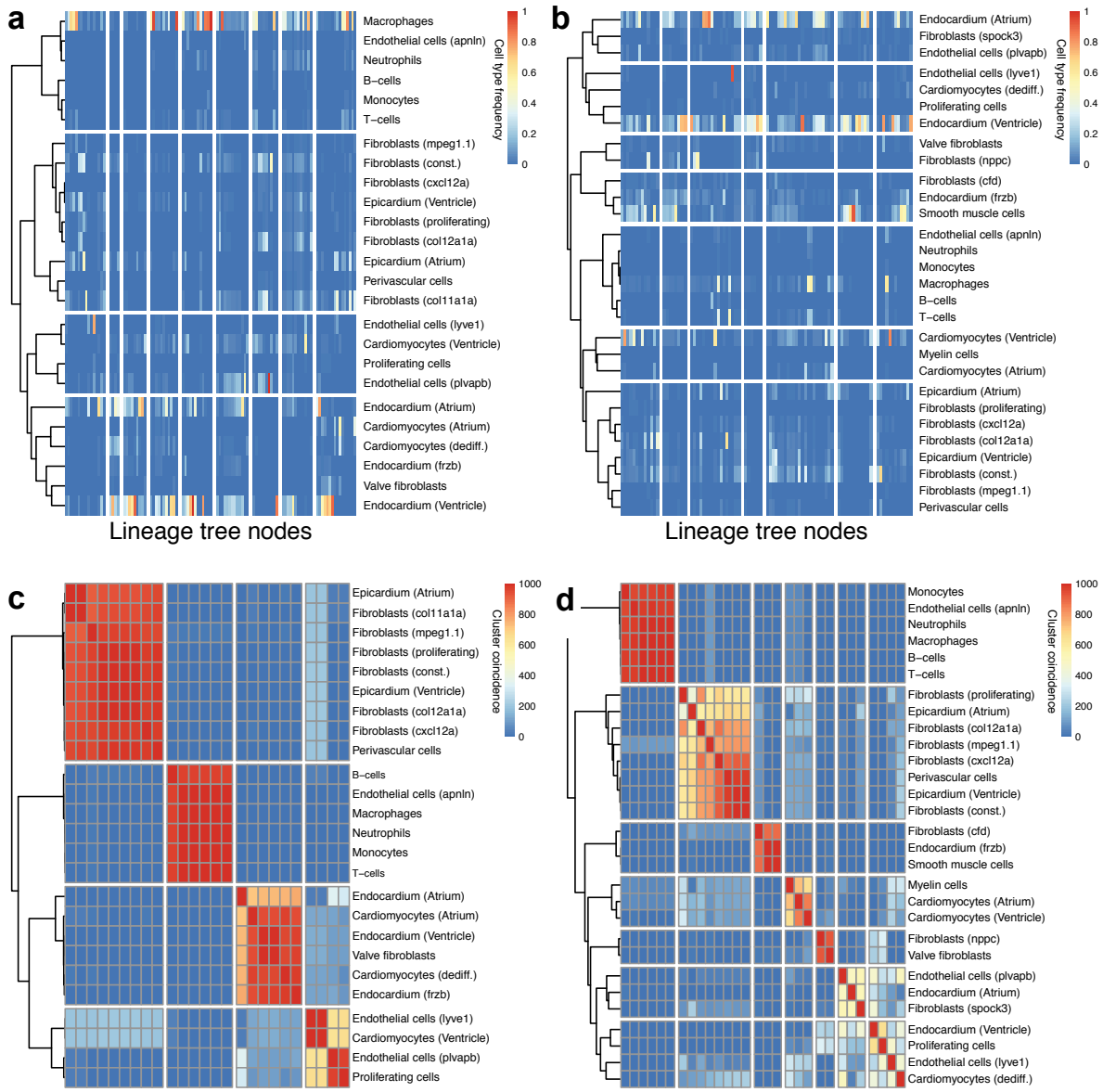
Additional Figure 9.



Additional Figure 9. Examples of lineage trees.

The heterogeneity of the tree topology is due to the stochastic nature of the barcoding events.

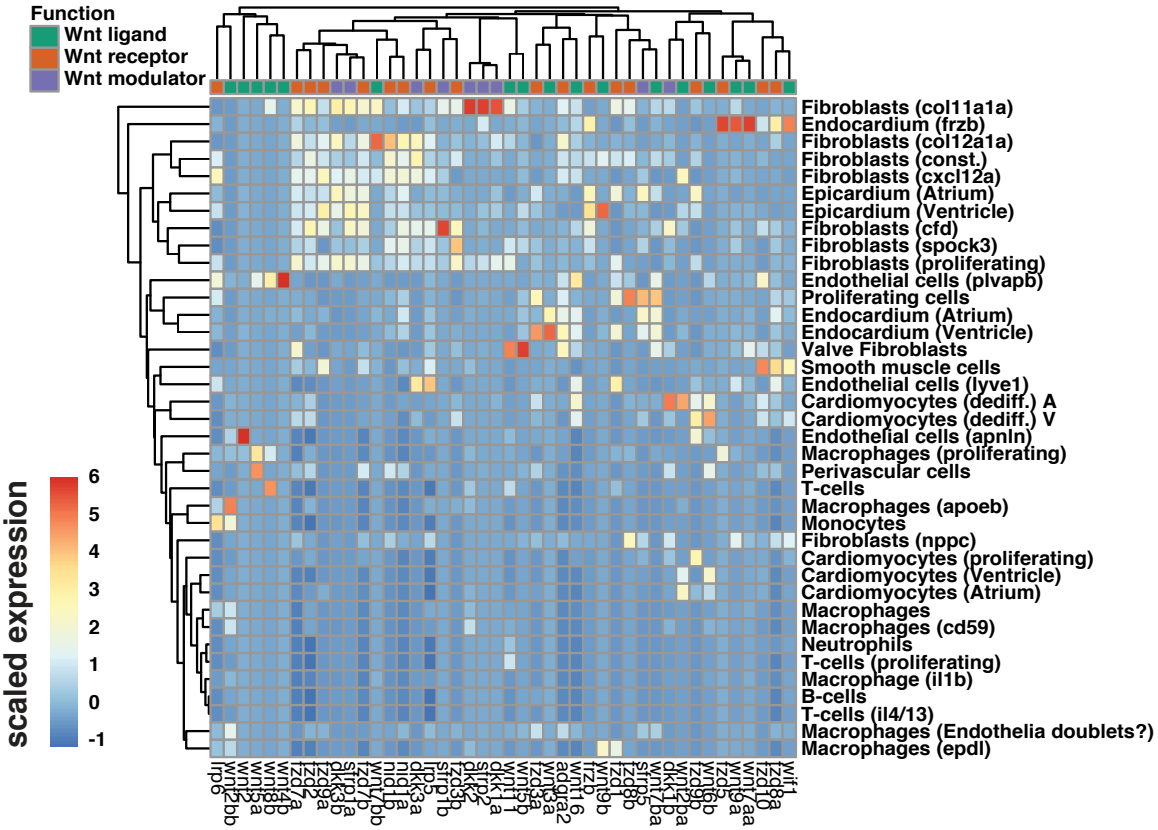
Additional Figure 10.



Additional Figure 10. Identification of lineage-related cell types at 3 dpi and 7 dpi.

a-b. Cell type fractions per node in 3 dpi (a) and 7 dpi (b) trees. **c-d**. Clustering at 3 dpi (c) and 7 dpi (d) is stable under 50% downsampling of trees.

Additional Figure 11.



Additional Figure 11. Expression of Wnt signaling factors across all cell type sub-clusters.

Wnt signaling factors are divied in Wnt-ligands, Wnt-receptors and Wnt-modulators. The selection of genes show all common canonical Wnt signaling factors.

Additional tables

Additional table 1.

Library statistics

Library name	Tissue/Sample type	Days post injury (dpi)	average nUMI per cell	average nGene per cell	number of cells	Chromium Chemistry	Wnt inhibition
L1	F1 larva	0	4143	731	11802	v2	none
L2	F1 larva	0	4076	747	13496		
L3	Larva	0	2825	743	5429		
L4a	Larva	0	2959	657	4211		
L4b	Larva	0	2589	634	2697		
L6	Larva	0	2529	568	2613		
L7	Larva	0	3476	753	13516		
L8	Larva	0	3341	756	18488		
T5	Telencephalon	0	1434	520	1133		
H5	Heathly heart	0	3924	816	1677		
P5	Pancreas primary islet and liver	0	7162	1046	3753		
T6	Telencephalon	0	1582	606	1017		
H6	Heathly heart	0	2695	591	3291		
P6	Pancreas primary islet and liver	0	6854	915	5101		
T7	Telencephalon	0	1696	658	4895		
H7	Heathly heart	0	4253	714	7280		
P7a	Pancreas primary islet only	0	10297	1122	3102		
P7b	Pancreas primary islet and liver	0	5944	781	8821		
H8a	Heathly heart, Atrium only	0	2286	608	5245		
H8v	Heathly heart, Ventricle only	0	2812	682	3728		
Hr1	Injured heart	7	4035	837	13027		
Hr2a	Injured heart	7	3247	725	10832		
Hr2b	Injured heart	7	3347	730	11183		
Hr3	Injured heart	30	3991	793	8575		
Hr4	Injured heart	30	3313	731	7367		
Hr6a	Injured heart, Atrium only	7	4656	938	2851		
Hr6v	Injured heart, Ventricle only	7	4145	897	7854		
Hr7a	Injured heart, Atrium only	7	3159	730	4332		
Hr7v	Injured heart, Ventricle only	7	3411	783	5511		
Hr8	Injured heart	7	2934	830	5604		
Hr9	Injured heart	7	4361	993	8829		
Hr10	Injured heart	3	1985	608	10063		
Hr11	Injured heart	3	1019	347	9887		
Hr12	Injured heart	3	3044	676	6869		
Hr13	Injured heart	7	1737	534	3835		
Hr14	Injured heart	7	2035	518	5449		
Hr15	Injured heart	7	1319	432	5207		
Hr19	Injured heart	30	2061	546	414		
Hr20	Injured heart	30	1910	533	419		
Hr21	Injured heart	30	2128	556	457		
Hr22	Injured heart	3	6364	1279	2140	v3	
Hr23	Injured heart	3	6942	1343	2689		
Hr24	Injured heart	3	5894	1092	4560		
Hr25	Injured heart	3	6164	979	1140		
Hr26	Injured heart	3	6873	1208	8150		
Hr27	Injured heart	3	6119	1132	5861		
Hr28	Injured heart	3	5619	1038	3739		DMSO
Hr29	Injured heart	3	5101	922	1829		IWR-1
Hr30	Injured heart	7	4703	972	3345		DMSO
Hr31	Injured heart	7	4680	693	1768		IWR-1
Hr32	Injured heart	7	6394	1169	1422		IWR-1
Hr33	Injured heart	7	5695	1057	1145		IWR-1
Hr34	Injured heart	3	8753	1571	5639		IWR-1
Hr35	Injured heart	3	8475	1625	7225		IWR-1

Additional table 2.

Cell clusters and differentially expressed genes for 5dpi larvae.

	Undifferentiated mesodermal cells	Skeletal muscle cells A	Neurons A (Brain)	Muscle cells (Pectoral fin)	Neuronal precursors A	Fibroblasts A	Neurons B (Proliferating)	Epidermal cells A	Erythrocytes A	Cone cells	Epithelial cells (Fin)	Skeletal muscle cells B
# cells	7057	4279	4164	2708	2569	2205	1945	1873	1718	1534	1517	1504
# of av. UMIs	2045	702	1185	1029	3127	2830	3675	5549	3278	1374	5449	15878
# of av. genes	594	328	491	368	847	705	878	960	579	457	1069	1218
top 50 genes	ifitm1	mylpfb	tubb5	hmgb2b	her4.2.1	tnw	hmgn2	cyt1l	hbae3	opn1mw1	and3	myhz1.1
2	cxcl12a	tnnt3b	tuba1c	hmgb2a	her4.1	vwde	hmgb2b	cyt1	hbbe1.1	opn1sw1	cldni	ttn.2
3	serpinh1b	actc1b	stmn1b	tma7	si:ch211-193l2.4	tgfbi	hmga1a	CABZ01059409.1	hbbe2	pde6h	apoeb	myhz2
4	serpine1	acta1b	gpm6ab	cnbpa	her4.2	si:ch211-251b21.1	si:ch211-193l2.7	krt17	hbbe1.3	gngt2b	and2	myhz1.2
5	pmp22a	ak1	zc4h2	cct5	si:ch211-193l2.3	col1a2	hmgb2a	krt4	hbae1.3	opn1sw2	pfn1	ttn.1
6	fstl3	tpma	elavl3	pabpc1a	her15.1.1	postnb	h2afvb	icn2	fth1a	gnb3b	tmsb1	pvalb1
7	rgs5a	pvalb2	nova2	rpl22l1	her15.1	tnmd	rrm2.1	krt97	cahz	rbp4l	cfl1l	tnni2a.4
8	twist1a	tnni2a.4	zgc:153426	ran	si:ch211-193l2.7	serpinf1	AL929185.2	im:7150988	blvrb	arr3b	si:ch1073-406l10.2	tnnc2
9	si:ch73-335l21.4	nme2b.2	rtn1a	eif3ea	her4.4	col1a1b	h3f3a	si:ch211-195b11.3	si:ch211-250q4.3	rcvrn2	CU693494.2	pvalb2
10	ms4a17a.8	mylz3	rtn1b	hmgn2	si:ch211-193l2.5	tmem119b	nusap1	mid1ip1a	hbae5	opn1lw2	epgn	ckmb
11	zgc:153704	myl1	gpm6aa	stmn1a	her12	fmoda	fabp7a	si:ch211-117m20.5	si:ch211-207c6.2	gnat2	epcam	atp2a1
12	CABZ01092746.1	mylpfa	tuba1a	eif3ba	her2	serpinh1b	hmgb1b	zgc:193505	hbae1.2	arr3a	and1	myhz1.3
13	krt8	ckma	ebf3a	setb	her4.3	dcn	hes6	zgc:111983	si:ch211-103n10.5	ckmt2a	aqp3a	mylz3
14	vim	tnni2b.2	gng3	seta	AL929185.2	cxcl18b	tubb2b	krt96	si:dkey-25o16.2	si:ch1073-469d17.2	si:rp71-77l1.1	si:ch73-367p23.2
15	mfap2	pvalb1	hmgb3a	ranbp1	si:ch73-21q5.7	si:ch211-286o17.1	stmn1a	CABZ01059413.1	prdx2	prph2a	si:dkey-248q15.3	mylpfa
16	si:dkey-12l12.1	mybphb	cd99l2	cirbpa	hmgb2b	twist1a	her15.1	cldne	tspo	si:ch211-81a5.8	cxl34b.11	tpma
17	cyr61	slc25a4	calr	hmgn7	hmgn2	pmp22a	pcna	si:dkey-248g15.3	zmp:0000001081	prph2b	rbp4	neb
18	klf2a	tnnt3a	elavl4	eif3f	id1	col5a1	si:ch211-288g17.3	FO904881.1	zgc:56095	guca1c	zgc:158343	ckma
19	jun	ckmb	ptmaa	hmga1a	hmga1a	jdp2b	si:ch211-222l21.1	lye	nt5c2l1	tmem237a	ptgs2a	myl1
20	hspb1	si:ch73-367p23.2	tmeff1b	eif5a2	hmgb1b	sparc	mki67	stard14	slc4a1a	ndrg1b	tagln2	slc25a4
21	ecrg4a	myl10	si:dkey-276j7.1	cct4	si:ch211-222l21.1	socs3b	seta	anxa1c	gpx1a	slc25a3a	eqfl6.1	atp2a1l
22	wnt11r	myl13	sox4a.1	zgc:56493	sox3	her6	cdk1	cldnb	BX936337.1	si:ch211-285j22.3	bcam	actn3a
23	fosab	si:dkey-16p21.8	sox11b	hnrnpa0b	hmgb2a	mfap2	khdrbs1a	si:ch211-173d10.1	mibp	si:ch73-28h20.1	zgc:153911	tnnt3b
24	igf2b	CELA1 (1 of many)	hmgb1b	tpma	h2afvb	cd82a	h2afva	si:ch1073-340i21.3	lmo2	pdcb	tmsb4x	si:ch211-255p10.3
25	krt18	aldoab	hsp90b1	psmb7	dla	rbp4	ran	anxa1b	alas2	gngt2a	zgc:86896	tnni2b.2
26	cebpd	ela2	sncb	atp5b	BX465834.1	col1a1a	si:ch73-1a9.3	si:ch211-207n23.2	hbz	irbp	KRT18 (1 of many).1	pkmb
27	si:ch211-251b21.1	pvalb4	ywhaq2	hnrnpabb	si:ch211-288g17.3	matn4	tubb4b	epcam	epb41b	ckbb	zgc:162509	actc1b
28	mycb	tnnc2	gng2	tnnt3b	lfng	zgc:162730	cirbpa	spint2	nmt1b	guk1b	krt5	mylpfb
29	rpz5	fabp1b.1	atf7ip	mibp2	rrm2.1	tmem176	cirbpb	caspb	tmem14ca	si:ch1073-303d10.1	krt17.1	actn3b
30	KRT18 (1 of many)	afp4.1	si:dkey-56m19.5	eif3d	hes6	sdc4	tuba1a	aep1	znfl2a	cnga3a	hbegfa	myoz1b
31	cyp2ad3	ela3l	si:ch211-288g17.3	eif2s2	elavl3	rps3a	snrpf	zgc:85975	mibp2	ppdpfa	thy1	ak1
32	sdc4	zgc:136461	hnrnpa0a	myl23	tubb2b	rpl10a	snrpb	KRT86	vdac2	si:ch211-207l14.1	cebpb	acta1b
33	junba	prss59.2	hp1bp3	ncl	sox11b	rps4x	lbr	pycard	blf	ncaldb	s100a10b	mybphb
34	zgc:162730	cpa5	hmp19	tubb2b	nusap1	rpl3	snrpe	si:ch211-95j8.2	CR381599.1	rgs9a	anxa1a	aldoab
35	tmem176	try	hnrnpaba	ak2	mdka	rps28	banf1	si:ch1073-190k2.1	tal1	eno1b	gstm.3	casq1a
36	hgd	ela2l	tubb2b	eif3ha	seta	rpl39	ascl1a	tmem176l.4	aqp1a.1	histh1l	ca6	nme2b.2
37	junbb	ctrb1	marcksb	qtpbp4	hmgn6	rps6	hnrnpabb	si:ch211-157c3.4	zgc:92066	nme2a	lgals3b	ryr3
38	mmp2	apoc1	marcksl1b	CABZ01020261.1	h3f3a	rpl7	hmgn6	cd9b	clic2	spock3	cotl1	ldb3b
39	ntd5	zgc:112160	hmgn6	tnni2a.4	hnrnpabb	rpl8	cldn5b	ch25h1.1	selj	si:dkey-17e16.15	rac2	ampd1
40	dcn	fabp10a	fez1	rbm8a	stmn1a	rplp1	si:ch211-156b7.4	zgc:110333	ap2m1a	si:ch1073-111c8.3	si:zfos-1011f11.1	si:dkey-16p21.8
41	hpdb	prss59.1	H2AFX (1 of many).5	snu13b	gas1b	rpl18	ddx39ab	zgc:175088	zgc:109934	igfbp7	igfbp7	myom2a
42	cd81a	fabp2	smarce1	h3f3a	fam60al.1	rpl26	sub1a	si:dkey-247k7.2	tmod4	gnb5b	cldn1	tcap
43	si:ch211-286o17.1	lgals2b	chd4a	nop10	si:ch73-1a9.3	rps18	sumo3b	soul2	sod1	unc119b	ets2	myom1a
44	col1a2	tmsb1	si:ch73-1a9.3	fkbp1aa	khdrbs1a	rps8a	snrpd1	cfl1l	tpd52l2b	si:dkey-65b13.13	ecrg4b	hspb9
45	socs3b	zgc:136930	mbd3b	h2afvb	hmgb1a	rps15a	marcksb	krt5	histh1l	syt5a	tp63	CR751608.2
46	ms4a17a.9	ctri	khdrbs1a	snrpb	nova2	RPL41	fabp3	agr1	bloc1s6	aqp9b	bzw1b	tmem38a
47	mxra8b	cox7a2a	ywhah	cct3	si:ch211-137a8.4	rpl32	ccna2	scel	glrx5	arl3l1	tmem238a	mt-co1
48	wu:fb55g09	mt-atp6	fscn1a	rpsa	hnrnpaba	ube2e1	ppiab	zgc:163030	gstm.1	cplx4a	si:dkey-102c8.3	tnnt3a
49	ackr3b	cox7c	h3f3b.1	rpl5a	si:ch211-156b7.4	rpl7a	six3b	sepw1	tuba8l4	arl13a	pak2a	pabpc4
50	jdp2b	krtt1c19e	sox11a	khdrbs1a	sb:cb81	rps19	mdka	anxa2a	anp32b	si:cabz01078386.1	myl12.1	smyd1a

	Chondrocytes A	Retinal cells A	Erythrocytes B	Epidermal cells B	Fibroblasts B (Fin)	Hepatocytes A	Satellite cells	Smooth muscle cells (Vasculature)	Intestinal cells A	Lens cells A	Fibroblasts C	Neuronal precursors B	Neurons C (Spinal cord)
# cells	1472	1415	1375	1355	1346	1335	1274	1253	1240	1195	1195	1137	1135
# of av. UMIs	2716	839	2805	9504	3096	2510	5337	2428	5334	636	3997	1850	1901
# of av. genes	718	382	401	1463	723	528	990	688	951	295	898	671	766
top 50 genes	otos	vsx1	hbae1.3	krt17	and2	apoa1b	hsp90aa1.1	acta2	fabp2	crygm2d2	fgl1	atp1a1b	elavl4
2	matn4	rs1a	hbbe1.2	CABZ01059409.1	hpdb	tfa	myog	tagln	chia.2	crygm2d1	fmoda	slc1a2b	sncb
3	lect1	neurod4	hbbe1.1	krt4	ucp1	apoa2	tmsb	BX088707.3	fabp1b.1	CR847505.1	si:ch211-287a12.9	fabp7a	snap25a
4	col2a1a	gnb3a	hbbe2	icn2	hgd	fabp10a	jam2a	aoc2	rbp2a	crygm2d20	vwde	glula	sncgb
5	col9a1a	syt5b	hbbe1.3	si:ch211-117m20.5	col1a1a	fetub	vwde	rgs5a	apoa1a	crygm2d3	col1a2	mdkb	gng3
6	col9a3	calb2b	hbae3	si:ch211-195b11.3	im:7150988	apom	unc45b	TPM1	gsta.1	crygm2d13	col1a1b	mfge8a	atp6v0cb
7	col9a2	hmp19	zmp:0000001081	cyt1l	tat	zgc:123103	npm1a	igfbp7	lgals2b	crygm2d14	tnw	zgc:165461	ywhaq2
8	cthrc1a	epb41a	hbae1	zgc:193505	ttc36	serpina1l	tuba8l3	myl9a	apoa4b.2	crygmx2	postnb	ptgdsb.2	vamp2
9	ucmab	gng13b	cahz	aep1	ucp3.1	apoc1	pabpc4	myl6	tm4sf4	crygm2d7	si:ch211-251b21.1	slc6a1b	si:dkey-7j14.5
10	tgfbi	snap25b	blvrb	cyt1	ecrg4a	uox	qkia	tpm4a	plac8.1	crygm2d5	rbp7b	cx43	stmn1b
11	fibina	rgs16	si:ch211-250g4.3	lye	CU929368.1	ces2	hey1	myl9b	apoa4b.1	crygm2d21	en1a	ptn	rnasekb
12	col11a1a	lin7a	fth1a	si:dkeyp-1h4.6	pcbd1	serpina1	rpl22l1	serpinh1b	afp4.1	crygm2d16	col1a1a	ptgdsb.1	stmn2a
13	mmp2	anp32e	prdx2	mid1ip1a	aqp3a	agxtb	zgc:92518	zgc:153867	zgc:158846	crygm2d6	sparc	atp1b4	stmn4
14	serpinh1b	mpp6b	si:dkey-25o16.2	si:ch211-207n23.2	and1	pklr	mibp2	eno3	s100a10a	crygm2d8	en1b	mdka	scg2b
15	sparc	si:ch73-1a9.3	gpx1a	zgc:111983	CABZ01092746.1	fgg	rplp2l	kctd12.2	cyp3a65	crygm2d4	col5a1	s1pr1	map1aa
16	twist1a	ppp1r1b	nt5c2l1	aqr1	c1qtnf5	hp.1	rps9	cfd	apoc2	si:ch211-255g12.6	mxra8b	si:ch211-66e2.5	atpv0e2
17	wfdc2	efna1b	zgc:56095	krtt1c19e	pah	ttr	eef1g	rbpms2b	apoea	crygm2d11	pdgfrl	her4.1	calm1b
18	cpxm1a	grb2b	hbae5	FO904881.1	ldhba	ambp	rps12	actb2	zgc:153968	crygn2	thbs4b	si:ch211-193l2.3	aldocb
19	epyc	h3f3b.1	si:ch211-207c6.2	anxa1b	qdpra	cfhl4	eef1b2	fhl2a	si:dkey-36i7.3	crybgx	fabp3	si:ch211-193l2.5	rtn1a
20	si:ch211-251b21.1	h3f3b.1.2	tspo	zgc:175088	fah	gc	RPL41	hspb1	si:dkeyp-73b11.8	crygm2d10	zgc:113307	si:ch211-251b21.1	cplx2l
21	col11a2	rsrp1	lmo2	si:dkey-247k7.2	gstz1	sid4	rpl11	ifitm1	gpx4a	cryba4	si:dkey-12l12.1	mt2	rtn1b
22	col2a1b	erh	hbz	fut9d	mdh1aa	rbp2b	rps3a	rbpms2a	sult6b1	crygm2d12	meox1	cspg5b	hmgb3a
23	pmp22a	CR383676.1	slc4a1a	si:dkeyp-51b9.3	bhmt	aldob	rpl17	cyp2ad3	cd36	cryba2a	entpd1	si:ch211-193l2.4	stxbp1a
24	sepp1a	gapdhs	BX936337.1	cldne	crabp2a	gamt	rpl21	sdc4	pck1	cryba1l1	mdka	crtac1a	gpm6ab
25	zgc:162730	vamp1	mibp	tmem176l.4	slc16a10	apoa4b.1	rpl10a	tmsb4x	zgc:172079.2	crybb1	twist1a	ndrg3a	snap25b
26	sdc4	hmgn6	mterf3	caspb	prrx1a	grhprb	rplp1	sparc	aqp8a.2	cryba2b	rbp4	her4.2.1	si:dkeyp-75h12.5
27	ntd5	nova2	histh1l	si:ch211-153b23.5	vcanb	si:dkeyp-73d8.9	rpl13a	si:dkey-7j14.6	gapdh	cryba1b	tcf15	efhd1	calm1a
28	cdh11	laptm4b	bloc1s6	soul2	msx2b	gapdh	rpl13	BX255935.5	si:ch211-133l5.7	crygm2d15	pmp22a	BX465834.1	syt1a
29	mfap2	oaz1b	selj	si:ch211-157c3.4	twist3	gpx4a	rpsa	fam212aa	si:dkey-60o16.5	crygm2d18	clec3ba	si:dkey-85k7.7	hmp19
30	mycb	btg1	faxdc2	pycard	hoxa13b	ITIH4 (1 of many)	rps24	pmp22a	faxdc2	crybb1l2	fbln7	id1	zgc:65894
31	tmem176	ptmaa	nmt1b	CABZ01059413.1	cbr1l	fgb	rpl23	hsp70.2	gsto2	crygm2d9	mfap2	her4.2	camk2n1a
32	rplp1	H2AFX (1 of many).5	znfl2a	krt5	msx1b	si:ch211-270n8.1	rpl37.1	ednraa	adh8b	crygm2d17	rpl7	slc3a2a	gapdhs
33	angptl1a	atp1b2a	tmod4	cd9b	prrx1b	crp	rpl7a	her6	krt92	crybb1l1	rpl8	hepacama	tmem59l
34	col11a1b	ppp1r10	mibp2	zgc:163030	mxra8b	agxta	rps15a	ptmaa	fbp1b	si:dkey-57a22.15	rplp1	her4.3	stx1b
35	col5a1	hspe1	CR381599.1	zgc:110333	glud1b	lect2l	rpl3	tmem176	cldn15la	fabp10a	cxcl18b	ppap2d	pcsk1nl
36	olfml3b	srsf3b	vdac2	stard14	kazald3	kng1	rps27a	col4a1	rtn4b	lim2.4	rps6	her4.4	atp6v1e1b
37	fmoda	pvalb2	fam53b	si:ch1073-340i21.3	serpine2	cdab	rpl23a	cxcl18b	stoml3b	crygmx	ube2e1	cd63	cbx1b
38	pdgfrl	ckbb	alas2	sepw1	msx3	ucp1	rps23	hsp70l	atp1a1a.4	apoa1b	myoc	ptprz1b	tmsb2
39	klf2a	actc1b	aqp1a.1	scel	zgc:92066	pnp4b	rps8a	mfap2	cox6b1	apoa2	rps3a	sox3	eno1a
40	fstl3	oaz1a	tmem14ca	epcam	abi3bpb	afp4.1	rps11	cd81a	gamt	fetub	ntd5	aldocb	nptna
41	timp2a	hnrpdl	ap2m1a	cfl1l	itgb1b.2	hpda	rpl26	gapdhs	aldob	pvalb1	krt8	slc1a3b	syt2a
42	mxra8b	hmgb1a	zgc:109934	si:dkey-87o1.2	fmoda	c9	rpl35a	tmem176l.1	ace2	prss59.1	rpl3	slc6a11b	tspan7b
43	ctsla	ip6k2a	CABZ01020261.1	si:ch211-95j8.2	ntd5	igfbp1a	rps7	dnajb1b	serpinb1l3	prss59.2	cd81a	syt11a	atp6v1g1
44	rpl32	si:dkey-177p2.6	ccni	si:dkey-201i2.1	twist1a	diabloa	rps19	snai1a	ada	actc1b	serpinh1b	slc38a3a	atp6ap2
45	tnmd	cct4	clic2	si:ch1073-190k2.1	gch1	fbp1b	rps14	tagln2	scp2a	try	RPL41	slc4a4a	dynll2a
46	rpl3	anp32a	zgc:92066	mgst1.2	dcn	serpinc1	ube2e1	sat1a.2	gsta.2	ckma	angptl4	gpm6bb	rab6bb
47	zgc:158343	atf7ip	dmtn	si:ch211-105c13.3	rps2	apoc2	rpl10	zgc:162730	tmem176l.2	pvalb2	angptl7	si:ch1073-303k11.2	eno2
48	rpl8	brdt	psph	si:ch211-173d10.1	sepp1a	scp2a	rps6	id3	serpinb1	mylz3	rpl32	bsg	upf3a
49	cyp26b1	atf4b	tpd52l2b	pglyrp5	prdx1	leg1.1	rpl27	tpi1b	ckmt1	tpma	rpl7a	itm2ba	cplx1
50	rps2	cct7	ISCU (1 of many)	plekhf1	cd82a	si:ch73-281k2.5	rps16	anxa13	BX908782.1	mylpfa	rpl26	tegt	atp6v1b2

Additional table 2. Cell clusters and differentially expressed genes for 5dpi larvae. (continued)

	Lymphocytes A	Keratinocytes A	Retinal cells B	Pigment cells (Xantophores)	Rod cells	Epidermal cells C	Fibroblasts D	Nephric duct cells	Neurons D (Brain)	Endothelial cells	Epithelial cells (Mucosa)	Lens cells B
# cells	1119	1091	949	947	932	856	815	803	790	767	737	696
# of av. UMIs	3434	6217	2306	2334	2509	4957	1892	3324	1650	3239	3372	3315
# of av. genes	746	1068	866	643	725	1028	629	897	604	859	796	426
top 50 genes	CR936442.2	si:dkey-183i3.5	rs1a	wu:fc46h12	rho	krt1-19d	krt94	aqp1a.1	stmn1b	fabp11a	si:ch73-306e8.2	crygmx
2	CR974440.2	zgc:136930	calb2b	gch2	gngt1	krt92	klf2a	podxl	pax10	vsg1	si:ch73-288o11.5	crybb1
3	pfn1	krt1c19e	rlbp1a	si:dkey-251i10.2	si:ch211-113d22.2	ponzr5	pmp22a	krt94	pax6a	dusp5	agr2	cyba4
4	CR936442.1	krt5	gnao1b	tmem130	gnat1	zgc:165423	si:ch211-39i22.1	cavin2b	ppdpfb	ramp2	cldnh	cryba1l1
5	arpc1b	CU693494.2	syt5b	qstp1	pde6g	krt97	vim	aoc2	ptmaa	igfbp7	icn	cryba1b
6	coro1a	krt17.1	cabp2a	slc22a7a	sagb	scel	rpz5	cxcl8b.1	si:ch73-1a9.3	srgn	muc5.1	cryba2b
7	ccr9a	cxl34b.11	snap25b	zgc:110339	rom1b	si:dkey-19b23.12	CABZ01051572.1	rhag	snap25b	ponzr1	cldnb	crybb1l1
8	cebpb	lgals1l1	glula	uraha	gngt2a	krt96	si:ch211-251b21.1	krt8	pax6b	cldn5b	KRT18 (1 of many).1	crygn2
9	rac2	cldni	vamp1	bscl2l	elovl4b	icn	mfap5	agrp	syt1a	tagln2	epcam	lim2.4
10	wasb	ptgdsb.1	vsx1	aqp3a	rs1a	si:ch211-157c3.4	KRT18 (1 of many)	jam2b	elavl3	clec14a	mal	si:ch211-255g12.6
11	laptm5	pfn1	gnb3a	akr1b1	si:ch73-28h20.1	icn2	jun	cxcl8b.3	hmgn6	etv2	si:ch1073-190k2.1	crybb1l2
12	srgn	rbp4	rgs16	KRT18 (1 of many)	saga	si:ch73-204p21.2	CABZ01092746.1	tmem88b	gpm6aa	myct1a	bik	crygm2d8
13	CD53	cfd	lin7a	paics	ckmt2a	agr1	krt8	tmem176	pvalb6	si:ch211-248e11.2	spint2	crygm2d13
14	arhgdig	col1a1b	mt2	mdh1aa	ppdpfa	anxa1b	si:ch73-335l21.4	rdh10a	sncb	sox7	zgc:101744	cryba2a
15	lcp1	postnb	nme2a	si:ch211-194k22.8	syt5b	FO904881.1	igfbp5b	tmem98	zc4h2	si:ch211-145b13.6	ponzr5	CR847505.1
16	qrap2a	tmsb1	nrn1b	prdx1	anp32e	krt4	si:dkey-261h17.1	cyp2ad3	h3f3d	ehd4	si:ch211-195b11.3	crygm2d10
17	qrap2b	cfl1l	si:ch1073-83n3.2	pax7b	fabp7a	anxa1c	osr1	cav1	tfap2a	gpr182	sepw1	crygm2d2
18	rgs13	apoeb	sh3gl2a	si:dkey-73n8.3	tulp1a	si:ch211-195b11.3	tmem176	rbp4	H2AFX (1 of many).5	plk2b	calm1b	crygm2d1
19	si:dkey-27i16.2	cxcl8a	glulb	zgc:113142	CABZ01102500.1	cldne	add3a	cfd	hmgb3a	si:dkeyp-97a10.2	tmsb1	crygm2d5
20	rps17	zgc:86896	fam107b	pmp22b	sypb	cd9b	sostdc1a	cd151	h3f3b.1	ecscr	si:dkey-193p11.2	crygm2d12
21	rplp2l	col1a2	ckbb	csrp2	neurod1	zgc:111983	htra1b	cavin1b	calm3a	fam174b	dnase1l4.1	crygm2d20
22	rplp1	si:ch1073-406l10.2	sncgb	phyhd1	crx	si:dkey-87o1.2	CR383676.1	cygb1	gpm6ab	tmem88a	si:dkey-87o1.2	crygm2d3
23	hsp90ab1	aep1	neurod4	zgc:92161	arl3l2	ppl	fosab	tagln2	qng3	si:ch211-33e4.2	tspan35	crygm2d21
24	cxcr4b	ecrg4b	eno1a	reep5	nr2e3	lye	pcolcea	krt15	stx1b	pecam1	gstp2	crygm2d16
25	rpl18a	epcam	bhlhe23	aldh9a1a.1	unc119b	anxa2a	rbp4	aldh1a2	marcksl1b	tmsb4x	cd9b	crygm2d11
26	ptpn6	rhbg	six3b	tspan36	zgc:109965	mid1ip1a	plpp3	hand2	ywhag2	cdh5	FO904881.1	grifin
27	nme2b.1	col11a1a	fabp7a	rtn1b	arl3l1	caspb	cyp2ad3	si:ch211-156j16.1	hmgn7	erg	mid1ip1a	crygm2d4
28	rpl8	epgn	epb41a	cyb561a3b	si:ch211-106m9.1	MAL (1 of many).1	mfap2	mfap5	cirbpb	she	f11r.1	mipb
29	rpl32	itgb1b.2	aldocb	dhrs13a.1	rbp4l	stard14	tagln2	ftr82	mdkb	si:ch211-156j16.1	tprg1	crygm2d7
30	rps7	aqp3a	atpv0e2	si:ch211-251b21.1	hmgn6	cldnb	cyr61	pmp22a	ckbb	krt18	cldn7b	crygmxl2
31	ube2e1	sparc	anp32e	pfn2	tmem244	si:ch211-95j8.2	crip1	lxn	scrt2	dll4	ucp3.1	crygm2d6
32	rpl34	col1a1a	cplx4a	aox5	ckbb	tmsb1	marcksl1a	mfap2	vamp2	clic2	cfl1l	crygm2d18
33	rps5	tmsb4x	acbd7	mlpha	rcvrn2	spint2	krt15	fabp11a	rsrp1	yrk	gnpnat1	crygm2d9
34	rpl7a	zgc:77439	efna1b	pts	atp2b1b	si:ch211-173d10.1	junba	sepp1a	stxbp1a	jam2a	si:ch1073-443f11.2	crygm2d14
35	zgc:56493	cldn1	atp1b2a	cax1	tmx3	cfl1l	sost	pcolcea	hmgn2	si:dkey-126g1.9	oclna	lim2.1
36	rps15a	capns1a	sypb	ucp3.1	arl13a	epcam	ccl25b	ehd1b	hmgb1a	cpn1.1	s100u	crybgx
37	rpl19	hbegfa	gapdhs	tmem243a	snap25b	si:ch73-347e22.8	ntd5	hspb1	gapdhs	crip2	si:dkey-248g15.3	crygm2d17
38	RPL41	spaca4l	ppp1r1b	pon1	si:ch73-1a9.3	si:ch1073-190k2.1	cd81a	rbpms2b	CR383676.1	eqtl7	fxyd1	crygm2d15
39	rpl18	KRT18 (1 of many).1	foxg1b	prdx6	TULP2	s100w	ifitm1	tmsb4x	tuba1c	podxl	hopx	mipa
40	rps6	tp63	calb2b	s100a10b	pde6d	ociad2	btg2	rpz5	hnrnpaba	msna	anxa5b	lim2.3
41	ppiaa	col17a1a	golga7ba	slc2a15b	msi1	s100v1	lxn	tspan18a	atp6v0cb	FO393393.1	lgals3b	crygn1
42	rpl3	mmp30	si:ch73-256g18.2	gpr143	neurod4	zgc:193505	zgc:162730	slc43a3b	khdrbs1a	afap1l1b	glrx	lim2.5
43	rps9	col2a1b	pcp4a	nrgna	h3f3b.1.2	nqo1	junbb	rnd3b	sncgb	grapa	si:ch73-347e22.8	endou2
44	rpl13	rac2	qabrr2a	rasd1	slc25a3a	anxa3a	si:ch211-105c13.3	cd81a	LINC00672	lmo2	elf3	bfsp2
45	rpl26	pak2a	efhd1	icn	atp1b2b	cyt1	wu:fb55g09	serpinh1b	h3f3b.1.1	tspan4b	si:ch73-234b20.5	capn3a
46	rpl10	kera	atp6v0cb	rspry1	gapdhs	prss60.2	krt18	sparc	zgc:153426	krt94	si:ch211-79k12.1	cryba1l2
47	si:ch211-105j21.9	col1	stx3a	prrtfdc1	atp6v0cb	tmem176l.4	jdp2b	anxa1a	dynll1	fli1a	btg1	si:ch73-111m19.2
48	si:ch211-161c3.6	si:dkey-102c8.3	mpp6b	gpd1b	foxg1b	apodb	plscr3b	si:dkey-7j14.6	hp1bp3	tuba8l3	rassf7b	cx23
49	rpl11	ca6	gng13b	pcbd1	guk1b	gsto2	si:ch211-286o17.1	si:ch211-286o17.1	calr	krt8	krt17.1	otc
50	rpl37.1	krt4	laptm4b	bace2	h3f3b.1	hsd17b12a	si:ch211-212k18.8	fn1b	srsf3b	tpm4a	tnks1bp1	crygm2f

	Pigment cells (Melanocytes)	Intestinal cells B (Exocrine)	Macrophages	Pigment cells (Iridophores)	Pancreas exocrine cells	Glial cells (Peripheral)	Skeletal muscle cells C	Radial glia A	Skeletal muscle cells D	Retinal pigment epithelial cells	Epidermal cells D
# cells	661	617	600	589	575	527	523	521	508	494	480
# of av. UMIs	5221	5475	3108	2479	7656	1224	10680	2778	23576	3269	8503
# of av. genes	959	1118	774	720	532	483	1222	774	1769	937	1329
top 50 genes	tyrp1b	fabp6	ccl34b.1	defb1	ctrb1	mbpa	pvalb7	epd	myhz1.1	rgrb	krt17
2	pmela	ctsl.1	MFAP4 (1 of many).4	apoda.1	try	mbpb	tnni2a.1	ggctb	myhz2	rlbp1b	cyt1
3	MREG	epdl2	si:dkey-5n18.1	lypc	prss59.1	cd59	tnni2b.1	soul5	pvalb1	rpe65a	si:ch211-195b11.3
4	slc45a2	lgals2b	mfap4	pnp4a	prss59.2	si:ch211-156j16.1	tpm2	fxyd1	ttn.2	mb	cyt1l
5	mtbl	anxa4	fcer1gl	IFI30 (1 of many)	ctrl	anxa13l	pvalb4	SLC3A2 (1 of many)	myhz1.2	fabp11b	zgc:111983
6	dct	ctsbb	lygl1	si:ch211-243a20.3	CELA1 (1 of many)	cldnk	TPM1	fabp11a	atp2a1	rgra	mid1ip1a
7	tspan36	s100a10a	ccl35.1	CABZ01067314.1	ela2l	si:ch211-132g1.3	myhz1.1.1	igfbp2a	pvalb2	dct	icn2
8	tyrp1a	BX908782.1	arpc1b	akap12a	zgc:112160	egr2b	myl1	mdka	ttn.1	irbp	krt4
9	icn	si:ch211-137i24.10	si:dkey-27i16.2	gpnmb	ela3l	nrgna	tnnt3a	cygb1	tnni2a.4	rdh5	CABZ01059409.1
10	s100a10b	grn1	havcr1	smim29	ela2	clic6	tpm4b	sepp1a	mylz3	tyrp1b	zgc:193505
11	anxa1a	si:dkeyp-73b11.8	tspan36	anxa4	zgc:136461	si:rp71-19m20.1	fhl2a	apof	tnnc2	lratb.2	stard14
12	mlpha	tmem176l.2	lgals2a	nme4	cpa5	asmtl	mylpfa	atp1a1a.2	ckmb	rbp5	cldne
13	gpr143	ifi30	pfn1	alx4a	si:dkey-14d8.6	cx27.5	mustn1b	rbp4	mylpfa	stra6	lye
14	gstp1	krt92	cebpb	zgc:110789	sycn.2	si:dkey-262k9.4	casq2	smox	atp2a1l	pmela	epcam
15	aldh9a1a.1	cldn15la	ctss2.2	ponzr1	c6ast4	efhd1	acta1a	wu:fj16a03	ckma	rbp4l	KRT96
16	rabl6b	si:dkey-36i7.3	lgals3bpb	si:ch211-256m1.8	zgc:92137	gldn	desma	tfpia	tpma	rdh8a	rhcga
17	CABZ01067314.1	slc10a2	coro1a	myo5b	apoda.2	pmp22b	atp2a1	slc3a2b	neb	proca1	fut9d
18	csrp2	wu:fb59d01	laptm5	tpd52l1	cpb1	cldn19	nme2b.2	atp1b4	slc25a4	tm6sf2	spint2
19	oca2	plac8.1	tmsb4x	CYSTM1	spink4	qki2	tpm3	lxn	si:ch73-367p23.2	pmelb	CABZ01059413.1
20	tspan10	cldn15lb	lgals9l1	si:dkey-21a6.5	si:ch211-240l19.6	myo1cb	myhz2	glulb	actn3a	ctsd	si:ch73-329n5.6
21	gstt1a	cldnc	atp6v0ca	emp3b	pla2g1b	pou3f1	ckmb	cx43	myl1	inpp5ka	anxa1b
22	zgc:110339	zgc:172079.2	si:ch211-212k18.7	tagln3b	amy2a	qpm6ab	lgals2a	lpl	tnnt3b	fqtbp1b.1	si:ch1073-190k2.1
23	slc22a2	serpinb1	vmp1	tfec	cpa4	nr4a2b	myhb	slc7a8a	actc1b	pttg1ipb	tdh
24	rtn1b	si:dkey-283b1.6	npc2	paics	zgc:92590	si:ch211-214p13.3	cav3	slc20a2	myhz1.3	tspan36	zgc:153911
25	smim29	tmigd1	ctss2.1	sytl2b	si:dkey-14d8.7	si:ch211-234p6.5	tcap	bsg	actn3b	lrata	rhbg
26	si:dkey-204f11.64	FO834808.2	ncf1	slc45a2	cel.2	vat1	tmem38a	slc13a4	myoz1b	si:ch211-237l4.6	si:ch211-157c3.4
27	zdhhc2	si:ch211-202h22.9	lcp1	fhl3a	cpa1	si:ch211-260e23.9	tnnt3b	flvcr2b	ryr3	slc16a1a	FO904881.1
28	slc2a15b	si:dkeyp-72e1.6	spi1a	CABZ01072077.1	sycn.3	sncgа	actc1b	tsc22d3	pkmb	fam213ab	zgc:110333
29	bace2	anxa2b	rnaset2l	phgdh	CELA1 (1 of many).6	fstl3	srl	bhmt	mylpfb	slc45a2	si:ch211-207n23.2
30	tfap2e	naga	CD53	cmtm6	pdia2	timd4	myl4	fstl1a	aldoab	msnb	cldnb
31	tmem243a	calml4a	spi1b	rgs4	fkbp11	zgc:153012	ckma	zic2a	hspb9	trpm1b	si:dkey-184p9.7
32	ppiaa	si:ch211-139a5.9	mpeg1.1	alx4b	c6ast3	depdc7	hspb11	zic3	ldb3b	aplp2	zgc:85975
33	slc39a10	eps8l3b	marco	cst14a.2	CELA1 (1 of many).3	pfn2	CABZ01069194.1	si:ch73-86n18.1	ampd1	slc4a5	si:ch1073-340l21.3
34	slc22a7a	cldn15a	c1qb	gmps	si:ch211-240l19.7	calm1a	mylz3	si:ch73-352p4.8	ak1	tyrp1a	chac1
35	pmpb	npc2	c1qc	ldhba	endou	phlda2	aldoab	slc6a22.1	fam13b	atp1b1a	pycard
36	si:dkey-21a6.5	si:dkey-28n18.9	si:ch211-147m6.1	s100v2	CELA1 (1 of many).2	sox10	tnnc2	cldn11a	pabpc4	slc16a8	tmem176l.4
37	mitfa	itpk1a	cxcr3.2	clec3ba	CELA1 (1 of many).4	rbm24a	AIMP1 (1 of many)	cyp1d1	mt-co1	trpm1a	si:ch211-95j8.2
38	slc38a11	CU682777.2	si:ch211-194m7.4	agtrap	si:ch211-255i20.3	ppp1r1c	si:ch73-367p23.2	cox4i2	eef2l2	tmem56b	anxa1c
39	AL935199.1	si:dkey-33i11.9	rac2	ppiaa	CELA1 (1 of many).1	CR383676.1	ttn.2	gpt2l	trim63a	oca2	ch25hl1.1
40	rpl12	ctsh	ctsz	lhfpl2b	si:ch73-44m9.5	histh1l	ak1	sat1a.2	si:ch211-255p10.3	rdh20	zgc:92242
41	pcbd1	dpep1	ptpn6	tcirg1a	aqp12	zgc:110340	hsp90aa1.1	atp1b1a	myom2a	pip4k2cb	zgc:163030
42	atp6v0ca	ezra	actb2	prps1a	nupr1	si:dkey-31g6.6	ttn.1	slc7a2	si:ch211-131k2.3	si:ch211-236l14.4	soul2
43	rasd1	serpinb1l3	si:ch211-102c2.4	uraha	si:dkey-266f7.9	grasp	ldb3b	slc38a3a	mybphb	cracr2ab	cd9b
44	gch2	prr15la	actb1	gpr143	cel.1	btg1	cavin4a	pdgfrl	si:dkey-51e6.1	igfbp7	scel
45	fabp3	tm4sf4	psap	FO393393.1	CELA1 (1 of many).5	plk2b	si:dkey-211f22.5	cyp3c1	myom1a	clic6	ppil1
46	prkar1b	anpepb	capgb	impdh1b	si:ch211-195b11.3	wasla	myoz1a	calm1a	smyd1a	slc13a4	cfl1l
47	RAB38 (1 of many)	si:ch211-133l5.7	scpep1	slc25a36a	cgref1	plekha1a	arpp21	slc25a5	si:dkey-16p21.8	slc7a2	si:dkey-201i2.1
48	pttg1ipb	CABZ01072043.1	irg1	pltp	SERP1	tspo	actn3b	pros1	hhatla	tspan10	plekhf1
49	spock3	si:ch73-234b20.5	tppp	lgals2a	gamt	fabp3	mybpha	osr1	si:ch73-366l1.5	calm1a	elovl7a
50	CABZ01088059.1	cox6a2	CU326347.1	zgc:110239	rtn1a	atf4b	txlnbb	itm2ba	klhl31	mfge8a	sepw1

Additional table 2. Cell clusters and differentially expressed genes for 5dpi larvae. (continued)

	Radial glia B	Chondrocytes B	Neurons E (Peripheral)	Neuromast cells	Epidermal precursors	Olfactory receptor cells	Corneal cells	Chondrocytes C	Smooth muscle B	Epithelial cells (Otic vesicle)	Hepatocytes B
# cells	415	405	386	369	350	346	344	341	339	317	297
# of av. UMIs	2109	1986	2162	5011	4841	2440	4093	4463	3471	4318	16602
# of av. genes	744	561	684	1188	1144	918	999	906	856	940	1687
top 50 genes	gfap	f13a1b	scn4ab	s100a11	si:dkey-33i11.4	pvalb5	scinla	matn1	tagln	otomp	tfa
2	stm	fgfbp2b	scn1ba	wu:fj16a03	si:dkey-193p11.2	s100z	gyg1b	mia	acta2	stm	apoa2
3	cldn7a	mia	si:ch211-134a4.4	tpm6	atp1b1b	zgc:173593	si:dkey-164f24.2	col2a1a	myl6	si:ch211-243g18.2	apoc1
4	vim	si:dkeyp-12a9.5	tuba8l3	zgc:109934	CABZ01021592.2	si:ch211-153b23.5	tkta	otos	myl9a	bricd5	apoa1b
5	endouc	serpinh1b	foxd3	FP016234.1	rgs5b	fabp10b	mfap2	col9a1a	csrp1b	aqr2	fabp10a
6	CLDN4 (1 of many).1	matn1	rgcc	si:dkey-33i11.4	cldnh	gstp1	twist1a	lect1	desmb	si:ch211-133n4.6	zgc:123103
7	zgc:165461	krt8	phlda2	zgc:198419	atp1a1a.2.1	zgc:109934	si:ch73-281n10.2	col11a2	TPM1	si:dkey-222f2.1	crp
8	pbxip1b	tgm2l	CLDN4 (1 of many).1	mafbb	foxi3b	cldnh	pmp22a	fgfbp2b	cald1b	si:ch211-80h18.1	serpina1l
9	hepacama	lect1	clic6	ponzr5	FO704910.4	txn	tubb6	col9a3	mylkb	si:ch211-152c2.3	serpina1
10	ptgdsb.2	hsp90aa1.2	sox10	pvalb8	si:ch73-359m17.9	sepw2b	cdk1	col9a2	myh11a	epcam	apoc1
11	atp1a1b	cbx7a	slc35c2	atp1b1b	clcn2c	chqa	aurkb	col11a1a	tpm2	si:dkey-16l2.17	apoa4b.1
12	igfbp7	gyg1a	gpm6ab	igfbp5a	cldnb	abhd2b	alpl	tnxba	myl9b	col9a2	fgg
13	fabp7a	gyg1b	krt8	si:dkey-159a18.1	dmrt2a	tspan2a	rplp1	epyc	fhl1a	cd9b	hp.1
14	hspb15	CR383676.1	KRT18 (1 of many)	cldnh	kcnj1a.1	TSTD1	hmgn2	si:dkeyp-12a9.5	lmod1b	CLDN4 (1 of many).1	pklr
15	cx43	cav2	hbegfb	capn8	slc12a10.2	ompb	tuba8l	ucmab	zgc:153867	col9a3	rbp2b
16	atp1b4	hspa4a	si:ch211-193l2.5	gcm2	si:ch211-270g19.5	capsla	mad2l1	prss35	zgc:110699	col9a1a	ambp
17	ptgdsb.1	pcolceb	alf3	prdx2	si:dkey-189h5.6	si:rp71-17i16.6	KRT18 (1 of many)	pcolceb	cxcl14	si:dkey-251i10.2	sid4
18	si:dkeyp-110a12.4	mcl1a	vim	ankfib	nqo1	vsig8a	kpna2	col2a1b	lxn	col2a1a	fetub
19	cav1	stip1	junba	dmrt2a	mal2	FP236545.1	hpdb	srpx	BX322618.1	serpine2	uox
20	wnt4b	pnrc2	pfn2	FO904881.1	COX5B (1 of many).1	htatip2	hmgb2a	itih1	ckba	sparc	ttr
21	hoxb8a	zfand2a	sb:cb81	cd9b	dnase1l4.1	TUBB4A (1 of many)	si:ch73-1a9.3	serpinh1b	cfd	fbln7	apom
22	ptn	aldob	serpinh1b	si:dkey-112a7.4	gsto1	tspan35	pitx2	si:dkey-6n6.1	rgs5a	wfdc2	fgb
23	ndrg3a	hsp70.3	marcksl1b	si:dkey-262k9.2	si:ch73-31d8.2	gng13b	nusap1	hapln1b	hand2	atp1b1a	gpx4a
24	mdkb	col9a1a	qkia	cfl1l	ndrg3a	cngа4	hmgn7	ddc6	rbpms2a	cldnb	atp4.1
25	ldlrad2	EPS8L2	tmem176	spint2	soul2	gnb1a	stmn1a	zgc:162730	cfl1	serpinh1b	cfhl4
26	slc3a2a	krt18	hepacama	si:dkey-87o1.2	atp5g3b	BX511021.2	ube2c	ecrg4a	aoc2	six1b	si:ch211-270n8.1
27	lpin1	clic2	fabp7a	tmem176l.4	sgk2a	adcyap1b	top2a	matn4	ppp1r12b	rassf7b	gc
28	cd81a	col9a3	fabp3	ctsl.1	s100v2	zgc:152698	rpl8	si:dkey-31g6.6	fxyd1	spint2	crp3
29	crip2	ppp1r1c	her15.1	slc10a7	epcam	zgc:101731	h2afvb	mcl1a	flna	zgc:92313	ITIH4 (1 of many)
30	BX465834.1	si:dkey-6n6.1	col4a2	tpte	atp5g1	si:ch73-335l21.2	rpl7a	arsia	ppp1r14aa	col11a1a	apoc2
31	hoxb7a	pdcd4b	her15.1.1	si:ch211-79k12.1	si:dkey-87o1.2	enkur	rps7	cpn1.1	BX927396.1	f11r.1	si:ch73-281k2.5
32	si:ch211-66e2.5	mknk2b	CR383676.1	tspan35	spint2	zgc:198419	rpl18a	cthrc1a	myocd	col9a1b	agxtb
33	sncga	oser1	cd81a	tmprss13a	si:ch1073-190k2.1	gnal	msna	tqm2l	pnp4a	sox10	c9
34	pdxka	hsp70.2	tmsb4x	cldne	sepw2b	prdx1	fstl3	gyg1a	si:ch211-137i24.10	cldn7b	lect2l
35	si:ch211-269l23.2	col9a2	rpz5	cox6b1	zgc:110333	si:dkey-148f10.4	hmga1a	papss2b	pdlim3a	fosl1a	apobb.1
36	crabp1b	pdcd6	fxyd1	si:ch1073-190k2.1	ndrg1a	scg2b	rpsa	cbx7a	tpm4a	hsp90ab1	grhprb
37	zgc:153704	hsp70l	col4a1	si:ch211-195b11.3	tspan35	si:dkey-71b5.7	rpl26	ptgdsa	cabp1a	hey1	gatm
38	id1	DNAJA4	btg1	tmsb1	echdc2	anxa5b	tpm4a	fqfr3	col6a1	hapln3	cyp2ad2
39	mt2	rsrp1	sparc	zgc:110333	prdx5	nxnl2	si:ch211-69g19.2	rlp26	actb2	matn4	pnp4b
40	ndrg2	dnajb1b	pmp22a	epcam	bik	si:ch211-10a23.2	fmoda	rcn3	plpp1a	CABZ01089828.1	gamt
41	aldocb	hsp90ab1	klf6a	cldnb	SYCE3	fstl5	rps15a	rpl7	il13ra2	sdc4	aldob
42	glula	ucmab	gulp1a	f11r.1	si:dkey-184p9.7	pcsk1nl	krt18	PLOD3	hlx1	cnn2	kng1
43	cav2	sdc4	sdc4	atp1a1a.2	slc25a5	bik	si:ch211-288g17.3	angptl1a	fhl2a	hsp90aa1.2	f2
44	mdka	anxa1a	si:ch211-137a8.4	fabp3	zgc:163083	CR391998.1	cks1b	rpl13	tmsb4x	ctgfa	serpina7
45	eno1b	ehd4	crip1	calm2a	FO904881.1	anxa13l	hgd	crtap	BX088707.3	foxp4	scp2a
46	cst3	sec61g	her9	vsnl1a	ppil1	gcnt7	sparc	eef1g	qucy1a3	atf3	c3a.1
47	sepp1a	col2a1a	krt18	zgc:158463	cldn7b	f11r.1	cfl1	slc35d1a	rbm24a	efhd1	leg1.1
48	tegt	ahsa1b	id1	tmem165	slc4a4b	FO904881.1	tuba8l4	pcolce2b	iscub	eya1	gapdh
49	slc4a4a	rplp2	nfkbiab	ifi30	slc25a3b	cplx2l	cks2	RPL41	gem	phlda2	shbg
50	sox19a	si:ch211-202a12.4	jun	plekhf1	si:ch211-195b11.3	nme5	zgc:153867	rps19	ntn5	hsp70.3	serpinc1

	Neutrophils	Osteoblasts	Skeletal muscle cells E (slow)	Oligodendrocytes	Epithelial cells (other)	Keratinocytes B	Hepatocytes C	Kidney cells	Chondrocytes D	Inner ear cells	Lymphocytes B
# cells	283	248	248	232	221	165	128	124	117	114	66
# of av. UMIs	3360	3079	5822	1928	4646	9231	17374	1789	4137	4251	4733
# of av. genes	770	774	1059	709	1086	1460	1757	576	853	1228	858
top 50 genes	lyz	col10a1a	myl10	cd59	si:dkey-205h13.2	CU693494.2	fabp10a	nit2	matn3a	s100l	si:dkey-21e2.15
2	npsn	si:dkey-22i16.3	myl13	plp1b	si:rp71-1c10.11	cxl34b.11	apoa1b	sepp1a	si:ch73-23l24.1	CU855484.1	si:dkey-21e2.16
3	srgn	ifitm5	tnnc1b	cldnk	stm	krt17.1	apoa2	acy3.2	col9a1b	CR925719.1	si:dkey-21e2.13
4	lect2l	spp1	tnni1d	si:ch211-156j16.1	si:dkey-4p15.5	cldni	fetub	aldob	zgc:113232	si:ch211-243g18.2	si:dkey-21e2.4
5	si:ch211-9d9.1	cd81b	ACTC1 (1 of many)	si:dkey-200l5.4	si:dkey-222f2.1	krt5	cbln5	TMEM27	zmp:0000000760	dnajc5b	si:dkey-21e2.10
6	fcer1gl	phex	tpm2	si:ch73-335m24.2	krt1-c5	ecrg4b	apom	pvalb9	epyc	si:dkeyp-110e4.11	si:ch211-212d10.1
7	mmp13a	panx3	tnnt2e	timd4	si:ch211-103f14.3	cldn1	f7i	atp1a1a	ucmab	rtn4rl2b	il4
8	arpc1b	sgms2	tpm3	mbpb	cbln20	zgc:86896	ttr	pdzk1ip1	cmn	gpx1b	MFAP4 (1 of many).1
9	pfn1	bgna	tnni1c	elovl1b	sox21a	epgn	ambp	tmem174	si:dkey-6n6.1	pvalb8	ccl34b.4
10	mpx	entpd5a	tnni4b.2	marcksl1b	moxd1l	CABZ01063789.1	si:ch211-113d11.5	hoga1	col2a1a	otofb	MFAP4 (1 of many).9
11	lta4h	si:dkey-32e6.6	myl4	cldn19	si:dkeyp-110c7.4	pfn1	uroc1	slc5a12	ta	gipc3	laptm5
12	scpp8	col5a2b	smyhc1	si:ch211-286c4.6	si:ch211-153b23.5	zgc:101810	gpx4a	pnp6	lect1	tmc2a	il13
13	coro1a	sod3b	myh7bb	si:ch211-133n4.4	zgc:158343	tp63	aqxtb	fbp1b	col9a3	calml4a	MFAP4 (1 of many).8
14	wasb	fpr1	smyhc2	mbpa	si:ch211-229d2.5	si:ch1073-406l10.2	si:ch73-15n24.1	gstr	chad	anxa5a	srgn
15	actb2	bmp8a	tnni1b	plekha1a	si:dkey-33m11.8	col11a1a	cbln17	si:ch211-139a5.9	matn3b	gb:eh456644	si:dkey-21e2.14
16	ncf1	col11a2	mybpc1	ppp1r14aa	phlda2	rac2	cbln11	slc22a7b.1	col11a1a	si:dkey-222f2.1	MFAP4 (1 of many).7
17	lcp1	pltp	acta1a	tuba8l3	fndc7a	rbp4	zgc:171534	slc26a1	si:dkey-234i14.3	si:ch211-284k5.2	pfn1
18	tmsb4x	PLA2G10	nme2b.2	cd82a	CABZ01021592.2	zgc:136930	a2ml	si:ch211-201h21.5	col9a2	twf2b	si:ch211-112c15.8
19	si:dkey-102g19.3	loxl4	hhatlb	tlcd1	CLDN4 (1 of many)	si:dkey-183i3.5	si:cabz01068815.1	slc25a5	gas2b	tspan13b	ccr9a
20	si:ch211-284o19.8	sparc	zgc:163073	fljl13639	tspan35	krtt1c19e	cfb	agxtb	si:dkey-12l12.1	CR792457.1	il11b
21	spi1b	si:dkey-12l12.1	myoz2b	qkl2	tmem176l.4	zgc:77439	miox	glud1b	shha	EMB	wasb
22	zgc:158446	si:ch211-76l23.4	tnnt1	si:ch211-132q1.3	irg1l	CU468917.1	hp.1	atp1a1a.4	sparc	baiap2l2b	scgn
23	il8r	serpinh1b	mustn1b	cd9b	cxcl14	KRT18 (1 of many).1	A2ML1 (1 of many).7	gpt2l	zgc:175280	ODF3L2 (1 of many)	grap2a
24	alox5ap	krt18	srl	cx27.5	cldnh	cfl1l	vtna	slc2a11l	si:ch211-243g18.2	si:dkey-229d2.7	MFAP4 (1 of many).10
25	CD53	mibp2	mybpc3	mag	selm	cxcl8a	si:ch211-212c13.8	atp5g3b	tgm2l	cabp2b	coro1a
26	si:ch73-359m17.5	si:ch211-170d8.8	atp2a2a	ugt8	si:ch211-152f23.5	si:ch211-117l17.6	si:dkey-86l18.10	ftcd	selm	si:ch211-218o21.4	grap2b
27	laptm5	pth1ra	pdlim5b	si:rp71-19m20.1	cldnb	si:dkey-91f15.1	comtd1	si:rp71-17i16.6	col11a2	si:ch211-193l2.10	arpc1b
28	si:ch1073-429i10.1	wif1	aldoab	lim2.2	tspan1	spaca4l	upp2	hrsp12	tpp3	nptna	capgb
29	ncorp1	tnmd	atp2a1	cdc42	pvalb8	entpd3	hal	gapdh	mmp23bb	enkur	FO203432.1
30	rac2	krt8	slc25a5	tmem125b	hspa5	ppp1r1b	si:ch211-288g17.4	ifi30	ehd2b	si:dkey-11c5.11	tnfrsf9b
31	capgb	col1a2	pvalb7	mpz	epcam	CR855337.1	vtnb	grhprb	SLC3A2 (1 of many)	si:ch211-172l8.4	CR847844.1
32	si:dkey-10p5.7	itga10	vmhc	slc44a1b	pdia3	ptgs2a	gc	gstt1a	rgcc	calm3a	rgs13
33	scinlb	ptgdsb.2	casq2	si:dkey-164f24.2	si:ch211-98n17.5	ca6	c8g	cox6a2	krt5	calm1b	CD53
34	lpar5a	bambib	pvalb4	si:dkey-204f11.64	manf	itgb1b.2	cp	epdl2	si:dkey-174i8.1	FO681289.1	BX927125.1
35	si:ch73-248e21.5	col1a1a	tnnt2c	ninj2	si:ch73-261i21.5	si:dkey-239h2.3	itih2	suclg1	golim4b	lrrc73	si:dkey-21e2.8
36	ch25h2	s100a10b	tnni2a	gpm6ab	prr15a	epcam	ITIH4 (1 of many)	gcshb	ctgfa	atp1b2b	si:ch211-67e16.3
37	si:dkey-19b23.14	emid1	desma	rhbdl3	si:ch211-152c2.3	itgb4	igfbp1a	ahcy	crabp2a	DYNLRB2	CABZ01053483.1
38	si:ch73-248e21.7	rpz5	hspb11	swap70b	net1	pak2a	cfhl5	cdh17	krt15	osbpl1a	si:ch211-102c2.4
39	cebp1	pmp22a	tnnt2d	tjp2b	si:ch211-243g18.2	aqp3a	slco1d1	mdh1aa	wfdc2	s100s	il10ra
40	cotl1	cdh11	rnf207b	olig2	cd9b	lgals1l1	apoba	slc20a1a	slc38a5b	pvalb9	cebpb
41	sepw1	id3	ryr1a	syngr2b	arg2	atp1b3a	serpinc1	atp5b	slc43a1a	daw1	CU570881.2
42	sult2st1	ndnf	tcap	fabp3	ndnfl	tmem238a	f7	slc13a3	igsf10	zgc:109934	rac2
43	lrrc15	timp2a	actc1a	cd226	soul2	tmsb1	plg	slc34a1a	slc3a2a	tekt3	arhgdig
44	mmp9	her6	ckmb	plp1a	cldn7b	hhla2a.1	si:ch1073-126c3.2	pdzk1	clu	rbm24a	si:dkey-53k12.2
45	mapre1a	tspan4a	gapdh	slmapb	slc1a3a	rhbg	cetp	slc2a12	slc43a2b	s100a1	sult5a1
46	ctss1	dlx5a	smpx	C7	si:ch211-195b11.3	sema3b	lcn15	aldh6a1	MMP23B	gltpb	si:dkey-12f6.5
47	arhgdig	anxa1a	ckma	crip3	KRT18 (1 of many)	mapk12a	pygl	zgc:136493	si:dkey-31g6.6	gfi1aa	zgc:153317
48	VSIR	rarga	fhl1a	zgc:153981	wu:fj05g07	si:rp71-77l1.1	agt	gamt	cdc42ep3	si:dkeyp-46h3.1	csf3a
49	glipr1a	igfbp3	MYH7 (1 of many)	hmgn3	rpz5	itgb1b.1	c3a.3	sord	haus7	lhfpl5a	tmsb4x
50	cxcr4b	hsp70.3	parvb	tmem136a	higd1a	col1a1b	cdab	cx32.3	si:dkeyp-12a9.5	atp2b1a	ptpn6

Additional table 3.

Cell clusters and differentially expressed genes for adult organs.

Cell clusters and differentially expressed genes for adult organs.

	Neurons (CNS)	Neurons (CNS)	Neurons (CNS)	Neurons (CNS)	Neurons (CNS)	Radial glia	Radial glia	Oligodendrocytes	Microglia	Microglia	Radial glia	Microglia	Unknown	Radial glia
# cells	1369	1094	986	813	596	517	314	255	233	224	214	207	155	68
# of av. UMIs	1043	825	1653	780	2359	3013	797	1624	863	1809	7576	3076	1780	2577
# of av. genes	499	402	765	385	1076	1086	435	710	363	538	761	821	673	1048
top 50 genes	CR383676.1	CR383676.1	nell2b	tubb5	slc6a1b	her4.2.1	fabp7a	si:ch211-286c4.6	cd74a	pfn1	epd	cd74a	icn	AL928901.1
2	TMEM269	tspan7b	adcyap1b	cd9912	gad2	her4.2	glula	si:ch211-132g1.3	cd74b	rgs13	ggctb	si:busm1-266f07.2	sepp1a	TUBB4A (1 of many)
3	slc32a1	adcyap1b	CR383676.1	tuba1c	slc32a1	her4.1	zgc:165461	cd82a	ccl34b.1	srgn	rbp4	cd74b	rpl12	enkur
4	gad2	chga	efna1b	tuba1a	gad1b	her4.4	acbd7	cd9b	apoc1	b2m	cp	ms4a17a.10	s100a10b	meig1
5	cacna2d2a	nptx1l	CABZ01073163.1	cirbpb	elavl4	her15.1	mt2	aplnrb	actb2	ccr9a	fxyd1	ccl34b.1	krt18	si:dkey-11c5.11
6	gad1b	asap1b	slc17a6a	stmn1b	sncgb	si:ch211-193l2.5	s100b	olig2	fcer1gl	rps17	wu:fj16a03	lgals3bpb	wu:fj16a03	pacrg
7	rtn1a	klf1aa	nrgna	elavl3	dlx5a	fabp7a	gstp1	aplnra	tmsb4x	RPL41	vtnb	zgc:158343	krt8	capsla
8	tnrc6c1	nell2b	gabra5	tubb4b	spock3	si:ch211-193l2.4	efhd1	si:busm1-57f23.1	si:busm1-266f07.2	rplp2	apoc1	tspan36	RPL41	s100a10b
9	nrxn2b	celf2	nrxn1a	tmsb	sst1.1	si:ch211-251b21.1	si:ch211-251b21.1	traf4a	pfn1	rplp2l	apoeb	ccl35.1	rps19	si:dkey-148f10.4
10	sdha	sltm	bhlhe22	khdrbs1a	vamp2	her4.3	ndrg3a	sema5a	lgals3bpb	sla2	soul5	havcr1	CNDP1	si:ch211-248e11.2
11	aldocb	serinc1	clstn1	h3f3d	srrm4	her15.1.1	zgc:153704	syt9b	lgals9l1	rpl37.1	qlulb	apoc1	rps27a	foxj1a
12	hnrnpa0b	clstn1	atp2a3	si:ch73-1a9.3	stxbp1a	msi1	lgals2a	tmem178b	arpc1b	rpl23	clec3ba	vmp1	atp1b1a	anxa2a
13	vdac2	appa	scg2b	ybx1	snap25a	fosab	qstr	zcchc24	rps2	rps2	c4	fcer1gl	fxyd1	CR293509.5
14	chd3	si:dkey-177p2.6	c1qtnf4	hmgn6	atp1a3a	zgc:165461	dap1b	sox10	AL928650.1	rplp1	slc13a4	ctss2.2	igfbp2a	spata18
15	slc6a1b	rnf10	kif1aa	si:ch211-222l21.1	si:dkey-175g6.2	si:ch211-193l2.3	mdka	qkia	rps2	rps29	SLC3A2 (1 of many)	AL928650.1	tfpia	spag6
16	atp1b1b	ip6k2b	gabrb3	hnrnpa0l.1	atp1b2a	her6	cx43	cd63	apoeb	ccl36.1	slc5a5	cxcr4b	rps7	si:ch211-243a20.3
17	chd5	cspg5a	vsnl1a	si:ch211-288g17.3	eno2	cx43	sepp1a	igsf9ba	gng10	rpl32	ccl25b	si:ch211-102c2.4	hsd11b2	si:dkey-27p23.3
18	eno1a	CABZ01073163.1	nrn1a	hmgb3a	nsfa	slc1a2b	ptn	si:ch73-335m24.2	rps15a	rpl36	mdka	slco2b1	rps27.1	tbata
19	atp6ap2	chd3	ak5	nova2	map1aa	cd82a	ddt	flj13639	eef1da	rpl18a	cygb1	lgals9l1	rplp2	zgc:174698.1
20	rnf10	sat1a.2	sncb	tubb2b	rnasekb	mdka	si:ch211-286b5.5	olig1	rps29	rpl36a	si:dkeyp-106c3.1	sdc4	rps2	lrrc51
21	ubc	appb	nptna	eef1g	atp6v1b2	cebpd	eno1b	itm2ba	ppiaa	rps27.1	sepp1a	laptm5	rpl36	ccdc40
22	ppp2r1ba	btg1	zbtb18	hnrnpaba	si:dkeyp-75h12.5	lipg	atp1a1b	bcan	rps17	rps15a	apof	mpeg1.1	rplp1	s100b
23	appa	epb41a	nptx1l	fabp3	rtn1a	dla	pdxka	rgcc	rpl19	CR753876.1	si:ch73-96n18.1	cebpb	rpl39	cfap52
24	eef1a1a	c1qtnf4	chga	gnb2l1	sv2a	si:ch211-66e2.5	sec61g	swap70b	lta4h	ccl38.6	bhmt	si:dkey-27i16.2	rpl26	txn
25	sltm	scg2b	gria2b	cirbpa	stx1b	atp1b4	ptgdsb.2	CT956002.2	rps28	rps28	pdgfrl	il10ra	ndrg1a	vim
26	serinc1	kif1b	atp6v0cb	rsrp1	atp6v0cb	s100b	slc1a3b	ctsf	rpl26	rpl35a	CNDP1	ccl35.2	rplp0	cfap126
27	nat8l	gas7a	syt1a	RPS17 (1 of many)	cdk5r2a	sox3	fabp3	cadm4	nme2b.1	rpl39	steap4	zgc:92066	rps26	armc4
28	calr	aplp1	atp1a3a	actb1	si:ch73-119p20.1	atp1a1b	prdx2	tuba8l3	rps14	cebpb	pltp	dnajb1b	rps15a	enpp6
29	atp1a3a	zgc:165603	bdnf	hnrnpabb	sypb	cxcl12a	gpx4a	gpm6ab	rpl35a	rpl34	SLC22A7 (1 of many)	zgc:103700	rpsa	cfap58
30	kif1aa	TMEM269	gpm6ab	atf4b	nrxn2b	gpm6bb	mgll	cxcl12a	rpl39	rpl35	flvcr2b	CD53	rpl29	cfap45
31	pabpc1b	egr4	camkvb	mt-co2	stmn2b	boc	sdcbp2	plxnb1a	rpl36	rsl24d1	colec12	ifi30	rps16	dnali1
32	qpia	qpm6ab	atp2b3b	ptmaa	atp6v1e1b	hs3st1l2	cox4i2	ptprea	rpl12	wasb	slc22a7b.1	c1qb	rpl36a	hspb15
33	kif1b	nrxn1a	rtn1a	ppdpfb	syt1a	tspan7	pnp6	atp1a1b	rpl32	CD53	tfpia	grna	rpl13	cebpd
34	nsfa	nrqna	celf2	tpt1	ywhaq2	id1	qabarapb	tspan18b	RPL41	rpl26	qpt2l	hsp70.3	rpl11	cetn4
35	epb41a	nptna	asap1b	hmgn7	pcsk1nl	ptgdsb.2	mid1ip1l	serpine2	rpl37.1	CR974440.2	col18a1a	csf1rb	rpl18a	daw1
36	rbm5	si:dkey-33c12.3	olfm1b	eif4a1b	nat8l	phgdh	s100v2	zgc:172122	anxa3b	zgc:162730	slc38a4	spi1a	rpl9	pih1d3
37	camk2n1a	zgc:65894	cdh8	sox4a.1	slc6a1a	ppap2d	slc1a2b	qria4a	rpl36a	rps21	iqfbp2a	zfand2a	rps10	ccdc17
38	cspg5a	olfm1b	camk2a	rpl24	cacna2d2a	si:ch1073-303k11.2	mdkb	myl1b	rps15	rps6	mmp2	hsp70l	rplp2l	PAMR1
39	slc3a2b	map1aa	qrin1a	rps15a	podxl2	si:ch211-286b5.5	cahz	tspan2a	rps19	rpl27	prosc	cmklr1	rpl8	BBOF1 (1 of many)
40	map1ab	lrrtm1	syt12	hmgb1b	hpcal4	glula	gapdhs	cst3	rps27a	CR936442.2	slc22a2	hspa4a	rps13	cfap77
41	spock3	tubb5	bsnb	rpl35	syngr3b	s1pr1	prdx6	scg3	rpl24	rps18	sort1a	hsp70.2	ctsla	cx43
42	rtn1b	rtn1a	cpe	rpl6	map4l	hsd17b3	mfge8a	vegfab	naca	rps11	si:dkey-8k3.2	pfn1	rps29	cd99
43	tob1a	tuba1c	atp2b2	rpl3	map1ab	tmem176l.1	luzp2	fabp7a	rpl23	ptprc	sfrp1b	BX677667.5	RPL37A	ak8
44	elmod1	sncb	fam43b	rpl28	si:dkey-7j14.5	ptn	sncga	rassf2a	rps8a	RPL37A	slc8a22.1	itgae.1	rpl37.1	fam228a
45	map1b	NA	pcsk1nl	rpl18	cplx2l	hmgn2	rtfdc1	TPM1	rpl34	rps24	vmo1b	apoeb	rps14	ccdc173
46	ywhag2	NA	pam	rplp0	npy	mfge8a	zgc:56493	picalma	RPL37A	dnajb1b	clu	mcl1a	faua	ccdc146
47	map1aa	NA	map2k1	atrx	sept5a	glulb	atp1b4	tmem50a	rpl17	rpl8	nmrk2	havcr2	rps28	gmnc
48	tmem59l	NA	cplx1	rpl30	dynll2a	mlc1	cox7c	slc1a1	rpl7a	rps26l	FO704741.1	rhbdd1	rps18	CR847939.1
49	aplp1	NA	zgc:165603	rps23	aldocb	crtac1a	itm2ba	atcaya	rpl22	rps7	slc6a9	gng10	rps4x	masp1
50	syngr3b	NA	sv2a	rps12	rab6bb	hepacama	ybx1	si:ch211-137a8.4	rpl35	dusp2	pros1	bzw1b	tpt1	EFCAB10

	Endothelial cells (Heart)	Erythrocytes	Endocardial cells	Cardiomyocytes (Atrium)	Cardiomyocytes (Ventricle)	Fibroblasts (Heart)	Erythrocytes	Cardiomyocytes (V)	Erythrocytes	T cells	Macrophages	Smooth muscle cells (Heart)	Endothelial cells (Heart)	Unknown
# cells	1565	1424	1279	941	754	715	687	656	624	580	547	439	262	236
# of av. UMIs	4.688	6.733	3.757	1.606	1.836	3.586	3.349	1.999	3.194	2.574	4.432	4.386	5.056	2.011
# of av. genes	1008	436	929	472	473	907	272	491	292	606	1054	1059	1074	407
top 50 genes	zgc:158343	hbaa1	si:ch73-86n18.1	tnnc1b	tnni4a	CR936442.2	HBZ (1 of many)	nme2b.2	hbaa1	pfn1	cd74a	rgs5a	vwf	rps19
2	ptgs2a	hbaa2	aqp8a.1	si:ch211-270g19.5	vmhcl	dpt	hbba2	eno3	hbba1	CD53	grn1	acta2	frzb	icn
3	ptgs2b	hbba1	fabp11a	tcap	fabp3	mdka	hbba1.1	gapdh	si:ch211-103n10.5	DNAJA4	ccl35.1	si:dkey-57k2.6	c2cd4a	rps26l
4	krt18	hbba2	f8	myh6	actc1a	mfap5	cahz	aldoaa	hbaa2	b2m	hmox1a	ALKBH3 (1 of many)	calm1a	rps15a
5	im:7152348	hbba1.1	mb	tnni1b	myl7	dcn	hbaa2	fabp3	hbba1.1	ccl34b.4	si:busm1-266f07.2	TPM1	krt15	rpl35a
6	id2b	si:ch211-103n10.5	spock3	aldoab	ak1	clu	slc4a1a	idh2	cahz	pdcd4b	cd74b	tagln	icn	rpl19
7	spock3	cahz	ramp2	mybphb	atp5h	CR974440.2	alas2	tnni4a	hbba2	CR936442.2	c1qb	ifih1	ecscr	rpl18a
8	mycb	si:dkey-25o16.2	epas1b	smtnl1	ndufa4	fn1b	zmp:0000001081	ckma	HBZ (1 of many)	hspa4a	c1qc	ctgfa	CR383676.1	rps8a
9	cxcl18b	igfbp1a	fam174b	cmlc1	acta1b	pmp22a	hbba1	ak1	zmp:0000001081	stip1	ctsd	si:busm1-57f23.1	esm1	rps18
10	si:ch73-335l21.4	slc4a1a	klf6a	nppa	hspb6	CR936442.1	si:ch211-103n10.5	myl7	si:dkey-25o16.2	cebpb	mafbb	lmod1b	tgfbi	RPL41
11	her6	blvrb	podxl	tnnc1a	cmlc1	sparc	CABZ01020261.1	ndufb7	si:ch211-250g4.3	coro1a	ctss2.1	myl6	ptmaa	rplp1
12	clic2	HBZ (1 of many)	si:ch211-145b13.6	tnnt2a	tnnt2a	phlda2	hbaa1	ldb3b	blvrb	capgb	fcer1gl	myh11a	krt8	rplp2l
13	fosl1a	si:dkey-16p21.8	cdh5	qapdh	qamt	CABZ01092746.1	si:dkey-25o16.2	ndufa4	igfbp1a	ccl36.1	lyql1	myl9a	f8	faua
14	zfp36l1b	zmp:0000001081	si:dkey-261h17.1	tpm4a	sepw1	rbp4	si:ch211-250g4.3	cox7a1	slc4a1a	mcl1a	rasgef1ba	krt17.1	si:ch211-153b23.5	rplp0
15	cfd	tspo	igfbp7	ckma	aldoaa	c6.1	nt5c2l1	actc1a	tspo	cd74a	marco	mdkb	s100a10b	SERP1
16	vcam1b	si:ch211-250q4.3	sid4	myl7	tnnc1a	sostdc1a	blvrb	vmhcl	qpx1a	tnfrsf9b	lqmn	flna	anxa2a	rpl39
17	dusp5	aqp1a.1	epgk1	cox5b2	mybpc3	c4	tspo	cmlc1	si:dkey-16p21.8	arpc1b	ncf1	vim	cpamd8	RPL37A
18	gpr182	qpx1a	si:ch211-248e11.2	idh2	cox17	gstm.3	si:dkey-16p21.8	atp5h	nt5c2l1	si:ch211-105j21.9	si:dkey-5n18.1	fhl1a	efemp2b	rps13
19	bzw1b	nt5c2l1	ecscr	atp5g1	cox5b2	htra1b	qpx1a	tnnt2a	alas2	ccr9a	atp6v0ca	KRT18 (1 of many)	cdh5	rps7
20	si:ch211-153b23.5	alas2	marcksl1b	acta1b	mdh2	col1a1b	igfbp1a	rbpms2b	aqp1a.1	erh	ctss2.2	tpm2	cav1	rps24
21	fosb	epb41b	tmem88b	mdh1aa	atp5g1	col1a2	fth1a	atp5g3b	fth1a	hmgn2	tspan36	cyr61	cyr61	rpl32
22	ctsla	mibp	cav1	mdh2	pdha1a	tmsb2	mibp	atp5e	epb41b	cxcr4b	grn2	si:dkey-164f24.2	krt4	rpl21
23	tmem88b	si:ch211-207c6.2	jun	mpc1	atp5ib	si:ch211-39i22.1	si:ch73-55i23.1	tpm4a	mibp	arhgdig	ctsz	c2cd4a	tgm2b	rps28
24	klf2a	klf17	lpl	calm3a	rbpms2b	col1a1a	gstm.1	USMG5	gstm.1	laptm5	ccl35.2	si:ch211-137i24.10	si:ch211-248e11.2	rps27a
25	marcksl1a	hbz	notch1b	atp5g3b	cox7a1	olfml3b	aqp1a.1	atp5j	hbz	scinlb	psap	slmapb	edn2	rps12
26	si:ch211-248e11.2	si:ch73-55i23.1	tuba8l3	atp5l	ppdpfa	col6a3	epb41b	mdh2	klf17	CR974440.2	atp6v1g1	si:ch211-1a19.3	ccdc80	s100a10b
27	serpinh1b	CR381599.1	tspan4b	actc1a	CR753876.1	c1qtnf5	dnase1l4.2	atp5o	CR381599.1	cd74b	si:dkey-27i16.2	LTBP4	rplp1	rpl37.1
28	akap12b	gstm.1	si:dkeyp-97a10.2	USMG5	mpc1	tmem176	rhd	tpi1b	hmgb2a	hspbp1	c1qa	tmsb4x	cd99	rpl26
29	igf2b	grb10a	atp5e	atp5e	ldhba	tcf21	rhag	cox6c	si:ch73-55i23.1	rgs13	vmp1	anxa1a	rps8a	rpl35
30	junba	zgc:56095	si:ch211-160j14.3	cox5aa	tpi1b	mmp2	pmaip1	atp5l	rhag	sla2	fgl2a	BX088707.3	junba	rpsa
31	grb10a	rhag	itm2ba	uqcrb	tnni1b	icn	hbz	uqcrb	si:ch211-207c6.2	calm1b	zgc:92066	alcama	rpl26	rps27.1
32	fosab	uros	akap12b	cox7a1	slc25a5	AEBP1 (1 of many)	CR381599.1	slc25a5	zgc:56095	ptprc	zgc:174904	htra1a	btbd6b	rpl29
33	glulb	slc20a1a	srgn	zgc:193541	cox6b1	s100a10b	si:ch211-197g15.10	atp5ib	slc20a1a	wasb	sqstm1	lpp	ehd2b	rpl7
34	hsp70.1	nmt1b	clic2	atp5ib	cox7b	soul5	prdx2	atp5b	tmod4	hsp90aa1.2	ms4a17a.10	PLA2G10	lmo2	rps2
35	sat1a.2	fth1a	kdrl	cox7c	atp5g3b	edil3a	mibp2	tnnc1a	metap2b	hspa8	rnfap4	actb2	clec14a	rps21
36	hsp70.2	gp1bb	anxa13	atp5f1	atp5d	vim	eif5	tnni1b	selj	fkbp4	scpep1	sh3bgrl2	rpl7a	rps26
37	btg2	si:ch211-197q15.10	zgc:64106	slc25a3b	eno3	f3b	uros	mb	tpd52l2b	cyb5a	si:ch211-212k18.7	myocd	rps9	rps14
38	krt8	selj	she	atp5o	cox7c	anxa2a	slc20a1a	cox7c	eif5	hspe1	zgc:109934	lsp1	net1	rps10
39	tmem88a	metap2b	krt94	slc25a5	uqcrb	hapln1a	zgc:56095	cox6b2	ncoa4	chchd2	slc30a1a	rasl11b	rpl3	rps16
40	aqp8a.1	si:ch211-207b24.4	s100a10b	atp5j	atp5g3a	mxra8b	metap2b	mdh1aa	prdx2	CR936442.1	hspe1	fhl3b	krt18	rpl13
41	f8	sgk2a	si:ch211-222l21.1	cox6b2	uqcrfs1	ckba	oaz1b	vdac2	NA	si:dkey-177p2.6	pfn1	ybx1	rpl30	rpl34
42	hsp70.3	alox5b.3	anxa2a	cox6c	zgc:193541	sepp1a	selj	cox7b	NA	si:ch211-165d12.4	lgals3bpb	cald1b	rpl10	rps5
43	ponzr1	gp9	icn	ybx1	tpm4a	rplp1	arl4aa	slc25a3b	NA	si:dkey-262k9.2	si:ch211-147m6.1	gch2	rpl32	rps27.2
44	ackr4b	si:ch211-227m13.1	krt8	atp5h	atpif1b	si:ch211-137i24.10	ucp3.1	atp5g1	NA	mknk2b	cxcr4b	prss35	hbegfa	rpl10
45	cebpd	slc43a3a	actb2	ak1	mdh1aa	actb2	bloc1s6	uqcrq	NA	srgn	CD53	CR855337.1	rps7	rpl23
46	lpl	eif5	CR383676.1	mb	atp5o	ALKBH3 (1 of many)	bnip3lb	atpif1b	NA	mhc1zba	grna	tpm4b	s100v2	rps15
47	junbb	sptb	rplp1	cox7b	cox6a2	si:ch73-380l3.2	NA	atp5c1	NA	wbp2	atp6v0e1	fhl2a	rpl13	rps17
48	pim1	hmgb2a	klf2b	pdlim5b	gapdh	pplb	NA	atp5d	NA	ncaldb	tmsb1	krt4	rpl10a	rpl24
49	wu:fb55g09	impdh2	her6	ldhba	mb	tbx18	NA	cox5b2	NA	rsrp1	arl6ip1	si:ch211-270g19.5	hsp90ab1	rpl18
50	ddx5	klf1	msna	eno3	ndufa4l	tmsb4x	NA	cox5aa	NA	CR383676.1	il10ra	pcolcea	bambib	rpl22

Additional table 3. Cell clusters and differentially expressed genes for adult organs. (continued)

	Smooth muscle cells (Heart)	Erythrocytes	Fibroblasts (Heart)	Cardiomyocytes (Atrium)	Apelin expressing cells	Monocytes	B cells	Endothelial cells (Blood vessel)	Fibroblasts (Heart)	Granulocytes	Pancreas exocrine cells A	Pancreas endocrine cells (delta)
# cells	229	225	199	199	168	165	140	91	87	36	1914	1408
# of av. UMIs	4.023	3.873	5.233	5.871	1.224	3.660	2.171	3.595	4.867	2.765	813	7341
# of av. genes	918	351	1116	1123	370	863	512	966	1039	663	247	1019
top 50 genes	fbln5	HBZ (1 of many)	gstm.3	nppb	hbegfb	id2a	zgc:194275	cxcl12b	tnfaip6	lect2l	ela2	cwc25
2	tagln	hbba1.1	s100a10a	hsp90aa1.1	pim2	arpc1b	mgst3b	igfbp7	fibinb	lyz	hbba1.1	sst1.2
3	itih1	hbba1	anxa2a	ttn.2	apln	epdl1	zgc:153659.1	rca2.2	zgc:153704	BX908782.3	prss59.1	nmba
4	wfdc2	hbaa1	krt15	nppa	si:ch211-195b11.3	zgc:92066	si:dkey-234i14.9	vsg1	angptl7	scpp8	prss59.2	BX120016.2
5	krt4	cahz	dkk3b	slc8a1a	admb	si:busm1-266f07.2	CU571382.1	CR381599.1	rspo1	npsn	amy2a	zgc:195023
6	cxcl12a	si:ch211-103n10.5	mmp2	si:ch211-131k2.3	fgl2a	cd74a	igic1s1	cldn5b	abi3bpb	pnp5a	cpb1	sst2
7	f3a	tspo	krt94	unc45b	gp1bb	cd74b	ighv1-4	sepp1a	ptx3a	mmp13a	cel.2	cldni
8	mdkb	blvrb	postnb	bag3	il11b	ccl35.1	cd74a	rgcc	igfbp5b	cebp1	ela2l	fabp6
9	acta2	fth1a	anxa1a	ttn.1	sele	mpeg1.1	CD37	plvap	krt4	il6r	CELA1 (1 of many).6	map2k6
10	myh11a	gpx1a	krt4	kcnh6a	rap1b	pfn1	cxcr4a	fabp11a	aif1l	cfl1l	ela3l	pdyn
11	ifitm1	si:dkey-25o16.2	si:ch211-105c13.3	myh6	rsrp1	sftpbb	iql3v5	pdqfbb	mqp	nccrp1	CELA1 (1 of many)	eqr4
12	ALKBH3 (1 of many)	nt5c2l1	si:ch211-286o17.1	ryr2b	itga2b	zgc:109934	cd74b	FO393393.1	cyp26b1	mpx	si:ch211-240l19.8	si:ch211-195b15.8
13	si:ch211-39i22.1	alas2	adh8a	xirp1	hsp70l	tmsb1	ccl35.1	aqtr2	carhsp1	si:ch1073-429i10.1	ctrb1	pitpnaa
14	krt17.1	si:ch211-250g4.3	mmel1	mt-co1	blf	zgc:152863	si:busm1-266f07.2	bcl6b	CR383676.1	srgn	si:ch211-240l19.7	ghrl
15	kng1	si:dkey-16p21.8	tuba1a	flnca	gcnt4a	syngr3b	cxcr4b	id1	csrp2	mmp9	c6ast3	nucb2b
16	lsp1	prdx2	tmsb1	atp2a2a	pfn1	ccl35.2	CR936442.2	glq2i	tuft1a	alox5ap	CELA1 (1 of many).1	c2cd4a
17	rpl7a	gstm.1	s100a10b	slc38a4	CR383676.1	cxcr4b	ccr9a	notchl	tgm2l	zmp:0000000924	zgc:92590	alcamb
18	cfd	si:ch1073-184j22.1	tmem98	lmo7a	hsp90aa1.2	spi1b	b2m	ca16b	gyg1b	scinlb	si:ch211-240l19.6	si:ch211-131k2.3
19	lmod1b	mibp	ckba	her15.1	tln1	grn1	DNAJA4	tmsb4x	irs2b	DNAJA4	try	slc7a7
20	vim	zgc:56095	krt8	tcap	srgn	actb2	hspa4a	col4a1	crip1	pfn1	cpa1	kctd12.2
21	cyp2ad2	slc4a1a	cxcl14	si:ch211-199o1.2	si:ch211-105j21.9	si:dkey-148a17.5	hsp90aa1.2	cavin1b	rplp1	antxr2a	si:dkey-14d8.6	scg3
22	rplp1	si:dkey-188i13.9	rbp4	ITGB1BP2	atp2b1a	zgc:173593	grap2b	ldb2a	gadd45bb	si:ch73-343l4.8	sycn.2	slc3a2a
23	rpl19	zmp:0000001081	igfbp2a	tnnc1b	Sep 12	si:dkey-33i11.9	hspbp1	actb2	vcanb	si:ch211-264f5.2	zgc:112160	ccl20b
24	si:dkey-14k9.2	tmod4	ppdpfa	CU929145.2	tuba8l	anxa3b	hspa5	CR383676.1	rpl22l1	si:dkey-19b23.14	zgc:136461	ifitm5
25	si:ch211-156j16.1	aqp1a.1	AEBP1 (1 of many)	myom1b	grasp	psap	pfn1	thy1	s100a10b	fcer1gl	cpa5	stmn1b
26	rps2	si:ch211-197g15.10	cygb2	zgc:86709	aldob	BX004816.4	CR974440.2	rgs4	ALKBH3 (1 of many)	ltb4r	ctrl	scg5
27	icn	si:ch73-55i23.1	CU468917.1	gnao1a	thbs1b	ctss2.2	CR383676.1	meox1	gsdf	plekhf1	CELA1 (1 of many).5	camk2n1a
28	rps18	eif5	icn	hspb8	tal1	hspa8	dusp2	cd81b	abhd3	tmsb1	serpina1	oaz1a
29	ptmaa	CR381599.1	fn1b	myom2a	dnajb1b	CD53	hspa8	cx43	timp4.3	si:ch1073-443f11.2	fabp10a	gulp1b
30	rps4x	isg15	hsd11b2	asb5b	bcl2l10	DNAJA4	irf4b	ch25h	qch2	cpa5	sycn.1	pax6b
31	rpl26	hbz	tuba1c	syne2b	si:ch211-214p13.9	zgc:136870	herpud1	fxyd6l	MYO1D	hspa8	apoc1	fxyd1
32	rps5	ucp3	phactr4a	syt1a	mpl	cyb561a3a	stip1	etv2	rpl19	CR848841.2	serpina1l	carhsp1
33	actb2	igfbp1a	flnb	popdc2	mcl1a	ctss2.1	pdcd4b	bsg	icn	si:dkey-102g19.3	tfa	mrap2a
34	rpl10a	impdh2	dclk1a	synpo2lb	myh9a	coro1a	dapp1	kcne4	bhmt	wasb	apoa1a	amd1
35	rpl18	rhaq	jam2b	ndrg4	si:ch211-33e4.2	hepacam2	rgs13	ptmaa	ITGB1BP2	si:ch211-284o19.8	fgq	zgc:194629
36	rps15a	slc20a1a	CU984579.1	bves	slc25a22	hspe1	pla2g15	rplp1	rpl26	hspbp1	apoa2	mdkb
37	cd248a	si:ch211-207b24.4	si:ch211-106h4.12	cavin1a	silk1	hsp90aa1.2	hnrnpa0l.1	ednraa	c1qtnf12	si:ch211-171h4.5	apoa1b	sesn1
38	rps11	guk1a	tcf21	ldb3a	pdcd4b	KRT18 (1 of many)	mcl1a	ackr3b	rpl7	hsp90aa1.2	CELA1 (1 of many).3	CABZ01020261.1
39	calm2a	rhd	mfap5	pygmb	hsp70.3	CR383676.1	si:ch211-150o23.2	tmsb1	ugdh	hspa4a	gapdh	TMEM179 (1 of many)
40	rpl7	klf1	lxn	her15.1.1	hacd3	ppiaa	si:ch211-195b13.1	CABZ01075125.1	serpinh1b	HACD4	pla2g1b	rab3db
41	rps16	s100a10a	tbx18	tnni1b	hsp70.2	gm2a	sh3bp5la	wu:fj16a03	rps7	zgc:77112	bhmt	serpinh1b
42	rpl35	epb41b	aldh1a2	myl7	cbx7a	stip1	hsp90b1	rgs5b	rps19	ncf1	leg1.1	pcsk1nl
43	rpl32	selj	c6.1	hspb11	hspa8	cotl1	ccl35.2	hopx	dcn	si:dkey-10p5.7	wu:fj16a03	jun
44	rps14	glulb	spaca4l	mylk3	brdt	atp6v1g1	rsrp1	actb1	fads2	adam8a	aldob	tmprss13b
45	myl6	rpp14	rps7	ahsa1a	afap1l1b	atp6v0ca	CD53	zgc:110340	thbs1b	chka	si:ch211-270n8.1	efr3ba
46	rps3a	si:dkey-188i13.7	slco2b1	trdn	DNAJA4	si:dkey-27i16.2	ncf1	grapa	TMEM269	arpc1b	crp3	ap3s2
47	rps8a	si:dkey-188i13.10	rps2	TMEM269	tmsb4x	hspa4a	chmp4ba	calm2a	rpl7a	BX908782.2	serpina7	nr4a1
48	rpl9	prkab1a	dcn	apcdd1l	hsp70.1	aif1l	cyba	rgs5a	rps15a	lta4h	adh8a	syt4
49	rpl13	ucp3.1	cd151	casz1	si:ch211-202a12.4	lgals2a	hsp70l	sox7	anxa2a	rnaset2	ccl39.2	slc26a3.2
50	ybx1	fam53b	sfrp1a	mybphb	si:ch211-222l21.1	BX897691.1	RPL41	pecam1	krt15	si:ch73-248e21.5	pglyrp6	srd5a2b

	Bi-hormonal endocrine cells	Pancreas endocrine cells (beta)	Smooth muscle cells B (Endocrine pancreas)	Pancreas endocrine cells (delta)	Hepatocytes	Pancreas exocrine cells A	Pancreas endocrine cells (alpha)	Smooth muscle cells A	sst1.1 cells	Hepatic stellate cells	Pancreas endocrine cells (delta)	Endothelial cells (Blood vessel)
# cells	1303	1254	1115	1067	1010	910	904	888	862	617	588	559
# of av. UMIs	1096	10351	3851	14226	23571	800	12552	1904	9405	7063	12552	4319
# of av. genes	258	1124	962	1762	1645	245	1155	687	1281	1218	1420	1025
top 50 genes	gcga	ins	myl6	sst2	apobb.1	ela2	gcgb	rasl12	sst1.1	gstm.3	sst2	fabp11a
2	gcgb	ppdpfb	csrp1b	map2k6	apoa2	prss59.1	gcga	CR381599.1	dkk3b	krt15	sst1.2	krt94
3	pnoca	ndufa4l2a	cygb1	egr4	apoa1b	cpb1	pnoca	angptl3	pyyb	podxl	si:ch211-195b15.8	zgc:195173
4	dkk3b	CABZ01020261.1	ccl19a.1	pdyn	fabp10a	amy2a	npb	kcne4	pdx1	anxa1a	pdyn	sele
5	ins	rgs5a	tagln	cldni	apoc1	prss59.2	dkk3b	si:ch211-251b21.1	srd5a2b	s100a10a	kctd12.2	cfd
6	sst1.1	scgn	acta2	sst1.2	c3a.1	cel.2	pcp4a	si:ch211-195b13.1	scgn	ptgdsb.2	fabp6	fam174b
7	prdx4	scg3	pcolcea	zgc:195023	si:ch211-270n8.1	ela2l	pyyb	si:ch211-270g19.5	rgs5a	s100a10b	map2k6	aqp8a.1
8	spcs1	ptn	ckba	slc3a2a	zgc:123103	ela3l	olfm1b	serpine1	scg3	anxa2a	zgc:195023	srgn
9	rnasekb	chga	myl9a	si:ch211-195b15.8	gpx4a	si:ch211-240l19.8	cfap58	frzb	ptn	cfd	cldni	krt8
10	mgst3b	pdx1	ifitm1	fabp6	ucp1	CELA1 (1 of many).6	ak5l	fxyd6l	pcsk1nl	krt4	pitpnaa	vsg1
11	eno1a	cxcl14	mylkb	pitpnaa	ccl39.2	CELA1 (1 of many)	scg3	agtr2	hmp19	ptgdsb.1	olfm2a	spock3
12	TMEM269	pim3	KRT18 (1 of many)	nucb2b	leg1.1	ctrb1	rgs5a	rasgef1ba	scg5	krt94	her15.1	ecscr
13	olfm1b	tagln3b	tpm2	c2cd4a	tfa	zgc:92590	scinlb	her4.1	chga	si:ch211-106h4.12	gulp1b	krt18
14	atp6ap2	scg2b	ccl25b	slc7a7	fgb	si:ch211-240l19.7	fev	gadd45ba	sox11b	krt17.1	mdkb	clec14a
15	zgc:158463	pcsk1nl	TPM1	alcamb	rbp2a	CELA1 (1 of many).1	scgn	slc20a1a	gck	ppdpfa	ifitm5	si:ch211-145b13.6
16	atp6v0cb	scg5	calm1a	dhrs1	serpina1l	si:ch211-240l19.6	slc7a2	BX322618.1	clu	icn	her15.1.1	zgc:172122
17	zgc:92818	rgs16	ALKBH3 (1 of many)	kctd12.2	crp3	c6ast3	chga	si:ch73-193i22.1	aldocb	aqp8a.1	slc7a7	pecam1
18	atp6v1e1b	tmsb1	myh11a	nmba	fgg	try	clu	tuft1a	cst3	mdka	rab3db	tmsb4x
19	scgn	rims2b	si:ch211-137i24.10	ifitm5	wu:fj16a03	sycn.2	cntnap2a	ifitm1	mtbl	fn1b	camk2n1a	si:ch211-79k12.1
20	aldocb	sncgb	tnfaip6	HSPA8 (1 of many)	ambp	cpa1	ptn	si:dkey-57k2.6	gfra1a	tmem88b	hepacam2	zgc:110340
21	ell2	tusc3	rbp4	efr3ba	serpina7	zgc:112160	ppdpfb	igfbp7	bdnf	aldh9a1a.1	mrap2a	plpp1a
22	calm1b	gck	aoc2	jun	hpx	si:dkey-14d8.6	nupr1	rgs4	pnoca	rbp4	ccl20b	rgcc
23	emc10	synqr2b	hsd11b2	egr2b	apoa1a	mt-co3	bckdk	iscub	scg2b	krt8	scg3	actb1
24	pcsk1nl	cpe	tpm4a	olfm2a	fetub	cpa5	slc43a2b	oaz2a	cpe	cxcl8b.3	slc28a3.2	cdh5
25	arl4ab	mnx1	rbm24a	oaz1a	hpda	ctrl	scg5	kctd12.2	g6pcb	lxn	nucb2b	lmo2
26	fxyd1	slc43a2b	fam107b	scg3	gstt1a	zgc:136461	CABZ01020261.1	junba	rtn4rl2b	eppk1	stmn1b	ramp2
27	qip	rtn4rl2b	glulb	carhsp1	igfbp1a	mt-co2	pcsk1nl	Sep 15	cacna2d2a	tmsb1	kcnk3b	grb10a
28	cst3	rnaseka	ppdpfa	gulp1b	serpina1	sycn.1	si:dkey-13a21.4	CABZ01068358.1	cpne4a	si:ch211-105c13.3	slc43a2a	cx43
29	clu	g6pcb	pmp22a	stmn1b	pnp4b	mt-nd2	pcxb	cebpd	SLC5A10	endouc	phkg2	si:ch211-248e11.2
30	scg5	c2cd4a	lxn	mrap2a	aldob	CELA1 (1 of many).3	slc1a4	sparc	cplx2l	slc43a3b	alcamb	igfbp7
31	nucb2b	clu	rims2a	mdkb	fga	mt-nd4	cplx2l	myl10	scg2a	ifitm1	syt10	krt15
32	ppdpfb	rnd1b	pdlim3a	map1ab	ahsg1	mt-nd3	nucb2b	calm2a	rims2a	cavin2b	ncf1	actb2
33	rgs5a	adm2a	rbp5	syt4	si:dkey-86l18.10	mt-nd1	camk2n1a	rbpms2b	ppdpfb	cavin1b	spock3	prdx1
34	nmba	prodha	flna	slc26a3.2	gapdh	mt-cyb	lgl2a	junbb	si:dkey-14k9.2	postnb	c2cd4a	ponzr1
35	tmsb1	si:ch211-255i3.4	zgc:153867	isl1	sid4	fabp10a	neurod1	sh3d21	tmsb1	c6.1	anxa11a	gpm6ab
36	nupr1	spon1b	cebpa	rab3db	ITIH4 (1 of many)	spink4	zgc:92818	TPM1	cited4a	actb2	srd5a2b	entpd8
37	c2cd4a	aldocb	jpt1b	sesn1	apoc2	mt-atp6	rgs16	CRIP2 (1 of many)	rgs16	aldh1a2	map1ab	jam2a
38	scg3	calm1b	mdkb	slmapb	leg1.2	apoc1	tusc3	BX927396.1	c2cd4a	s100v2	carhsp1	si:dkeyp-97a10.2
39	CABZ01020261.1	rab3c	cfl1	pcsk1nl	apoeb	apoa2	fam46c	BX255935.5	meis1a	cotl1	map3k15	neu3.5
40	HSPA8 (1 of many)	SLC5A10	ccng2	elovl1a	cfhl4	tfa	rnasekb	mcl1b	alcama	mmp2	BX548256.1	anxa13
41	egr4	rims2a	cxcl14	camk2n1a	ahsg2	apoa1b	calm1b	fam107b	wif1	mfap5	gck	erg
42	NA	ca7	tsc22d3	pax6b	CABZ01011067.1	leg1.1	rims2a	pgfb	slc1a4	gstm.1	cdx1b	id1
43	NA	ctsd	myl9b	rtn1b	rnasel2	wu:fj16a03	sybu	si:dkey-12l12.1	gpt2l	si:ch211-198c19.3	fev	cd81b
44	NA	uts1	COLEC10	ccl20b	uox	serpina1l	elovl1a	aoc2	KCNIP4	FAM84B (1 of many)	hopx	myct1a
45	NA	hspe1	SBSPON	ckbb	afp4.1	apoa1a	zgc:195356	ndufa4l2a	ca7	cav1	aebp1	ldb2a
46	NA	arl4ab	lmod1b	hist2h2l	cbln8	aldob	gpd1b	twist2	spon1b	c4	slc3a2a	ybx1
47	NA	rnasekb	si:ch211-270g19.5	nr4a1	adh8a	si:ch211-270n8.1	egr4	ogn	pfkfb4b	olfml3b	rasd1	she
48	NA	camk2n1a	c6.1	serpinh1b	pklr	gpx4a	FO904881.1	cd248a	ndufa4l2a	aoc2	slmapb	si:ch211-1a19.3
49	NA	olfm1b	mycb	TMEM179 (1 of many)	cp	zgc:123103	eno1a	rac1a	emc10	AEBP1 (1 of many)	syt4	sox7
50	NA	tp53inp1	fhl2a	ndrg3b	fbp1b	serpina1	adm2a	tmem119b	CABZ01067314.1	frzb	amd1	si:dkey-126g1.9

Additional table 3. Cell clusters and differentially expressed genes for adult organs. (continued)

	Pancreas exocrine cells B	Fibroblasts (Endoderm)	Pancreas duct epithelial cells	T cells	Smooth muscle cells A	Hepatocytes	Pancreas endocrine cells (beta)	Endothelial cells (Blood vessel)	Pancreas exocrine cells B	Pancreas duct epithelial cells	Endothelial cells (lymph)	Endothelial cells (Blood vessel)
# cells	519	498	494	470	467	426	373	349	324	315	302	269
# of av. UMIs	14608	3855	4943	2138	3931	2723	18910	1811	9638	5236	5150	3314
# of av. genes	553	880	969	554	1098	482	1620	652	478	1061	1127	982
top 50 genes	spink4	ccl25b	epcam	CR936442.2	acta2	tfa	ins	plvap	cpa5	zgc:193726	CU929150.1	si:ch211-237l4.6
2	zgc:112160	cxcl12a	lgals2b	pfn1	cxcl12b	aldob	ppdpfb	vsq1	ctrl	BX855618.1	sepp1a	cldn5b
3	try	col1a2	cfl1l	CD53	myl9a	apoa1b	ndufa4l2a	rgcc	CELA1 (1 of many).3	si:dkey-96g2.1	thy1	si:ch211-156j16.1
4	ctrl	col1a1b	si:ch211-153b23.5	hspa4a	taqln	apoa1a	mnx1	rca2.2	zgc:112160	lgals2b	lgmn	clec14a
5	CELA1 (1 of many).3	sparc	anxa4	stip1	si:ch211-76l23.4	serpina1l	rrad	aqp8a.1	try	her15.1	mrc1a	sox7
6	ela2l	col1a1a	tm4sf4	coro1a	cald1b	serpina1	ptn	plpp3	si:ch211-240l19.6	star	akap12a	efemp2b
7	cpa1	dcn	si:ch211-137i24.10	pdcd4b	BX927396.1	zgc:123103	pcp4a	vwa1	ctrb1	cd9b	arl4aa	ecscr
8	cpa5	AEBP1 (1 of many)	phlda2	arpc1b	TPM1	fetub	pdx1	ehd4	zgc:92590	SERP1	fabp11a	cfd
9	ctrb1	si:ch211-198c19.3	cldnc	DNAJA4	si:ch211-198c19.3	si:ch211-270n8.1	gpt2l	cd81b	CELA1 (1 of many)	cldn15a	lyve1a	jam2a
10	zgc:136461	mfap5	si:ch73-359m17.9	rgs13	lmod1b	apoa2	cxcl14	grb10a	si:ch211-240l19.8	slc4a4b	cdh6	fkbp1b
11	ela3l	ALKBH3 (1 of many)	si:dkey-96g2.1	wasb	cav2	fabp10a	tmsb1	mylz3	ela3l	cdh1	ramp2	klf2a
12	CELA1 (1 of many)	CABZ01092746.1	elf3	ccl36.1	myh11a	leg1.1	gng13b	exoc3l2a	cel.2	gmnn	clec14a	fabp11a
13	zgc:92137	pcolcea	hbegfa	cd74a	crip1	fgg	scgn	ucp3.1	cpb1	efna1b	gpr182	pecam1
14	si:ch211-240l19.7	c1qtnf5	cnn2	hspe1	fhl2a	pnp4b	rgs16	nrn1a	prss59.1	si:ch211-39i22.1	stab1	srgn
15	pglyrp6	actb2	cldn7b	ccl38.6	zgc:153867	pklr	chga	flt1	prss59.2	epcam	vwf	acvrl1
16	amy2a	s100a10b	cldn15a	CR974440.2	myl9b	adh8a	tmeff2a	si:ch211-145b13.6	sycn.2	zgc:198329	dab2	hbl2
17	cel.2	anxa1a	gpx1b	laptm5	BX908750.1	apoc1	cst3	kdrl	cpa1	eif4ebp1	cxcl12a	si:ch211-248e11.2
18	zgc:92590	icn	spint2	cd74b	rbpms2b	hpda	spon1b	tmsb1	si:dkey-14d8.7	her15.1.1	krt15	tmsb4x
19	pdia2	tagln	zgc:64022	cebpb	calm2a	gapdh	abhd15a	sec63	pla2g1b	cldnc	ackr3b	med13a
20	cpb1	col8a2	zgc:158343	ccr9a	prss35	rbp2a	wif1	igfbp7	si:dkey-14d8.6	cftr	rplp2l	rgcc
21	si:ch211-240l19.8	ccdc3b	f11r.1	b2m	tpm2	gpx4a	cbsb	spry4	endou	cx30.3	ecscr	krt94
22	si:dkey-14d8.7	tmsb4x	fam213b	chchd2	ptmaa	wu:fj16a03	scg3	epas1b	spink4	anxa4	sele	adgrl4
23	prss59.2	edil3a	ftr82	sla2	tpm4a	ahcy	slc43a2b	itm2ba	pdia2	spint2	ctsla	si:dkey-165a24.9
24	prss59.1	ifitm1	krt18	cyb5a	pkib	uox	neurod1	srgn	c6ast3	f3b	rab11bb	krt4
25	CELA1 (1 of many).1	c6.1	cd9b	rac2	actb2	serpina7	c2cd4a	cx43	amy2a	krt94	anxa2a	cdh5
26	si:ch211-240l19.6	RPL41	rassf7b	hspa8	zgc:162939	gstt1a	rab3c	ponzr1	ela2l	elf3	rpl39	si:ch211-145b13.6
27	pla2g1b	dpt	cdh1	si:ch211-102c2.4	cfl1	fgb	si:dkey-106c17.3	podxl	sycn.1	tm4sf4	etv2	si:ch211-1a19.3
28	si:dkey-14d8.6	zgc:153738	stard14	cxcr4b	bokb	fbp1b	tagln3b	jam2a	cpa2	icn	tmsb4x	dll4
29	sycn.2	ybx1	slc5a1	erh	si:ch211-10a23.2	hpx	sncgb	FO393393.1	cpa4	basp1	ponzr1	ptmaa
30	c6ast3	anxa2a	cdh17	ccng1	tpm4b	mdh1aa	scg2b	cdh5	CELA1 (1 of many).2	tmsb4x	glg2d	igfbp7
31	ela2	rpl39	btg1	wbp2	tmsb4x	ambp	mid1ip1b	ackr3b	si:ch211-195b11.3	entpd8	rplp1	ca16b
32	endou	sepw1	zgc:92313	tnfrsf9b	wfdc2	eef1da	si:dkey-153k10.9	smox	dnase1	si:ch211-137i24.10	anxa1a	si:dkeyp-97a10.2
33	cpa4	olfml3b	ccl19b	hspbp1	ifitm1	cfhl4	gck	anxa13	si:ch211-240l19.7	nr5a2	clic2	podxl
34	cpa2	rbp4	zgc:198329	fkbp4	rasl12	eno3	scg5	si:dkeyp-97a10.2	c6ast4	zgc:92066	rps19	rca2.2
35	CELA1 (1 of many).2	prelp	atp1b1a	mknk2b	krt17.1	crp3	copz2	msna	CELA1 (1 of many).1	aqp8a.1	igf2a	ponzr1
36	c6ast4	col6a1	actb1	grap2b	BX322618.1	ucp1	cntnap2a	adgrl4	ela2	ANPEP (1 of many).2	cd63	krt8
37	si:ch211-195b11.3	lgals2a	atf3	hsp90aa1.2	Sep 15	bhmt	tspan7b	fam174b	zgc:92137	ccl19a.1	CT030188.1	gpm6ab
38	dnase1	sfrp1b	tspan35	si:busm1-266f07.2	sost	ahsg2	pou2f2a.1	cfd	zgc:136461	clic2	marcksl1a	tmsb1
39	sycn.1	krt8	si:ch211-222i21.1	zdhhc18b	twist1a	apoc2	pcdh10a	nfe2l2a	CT027638.1	si:ch211-105c13.3	loxl1	id1
40	CELA1 (1 of many).6	ppib	icn	ahsa1b	irf2bpl	ccl39.2	foxa3	si:dkey-261h17.1	pglyrp6	fosl1a	rpl35	apoa1b
41	CT027638.1	KRT18 (1 of many)	si:ch73-335l21.4	hspd1	sparc	si:dkey-86l18.10	clu	fkbp1aa	CELA1 (1 of many).5	cldn7b	rpl18a	anxa2a
42	aqp12	rps28	ctqfa	mcl1a	slc20a1a	fqa	aclya	mhc1zba	aqp12	rassf7b	si:ch211-248e11.2	s100a10b
43	CELA1 (1 of many).5	pmp22a	actb2	rhbdd1	myl6	sid4	peli1b	atp1b1a	si:ch211-255i20.3	cnn2	rps6	hspa12b
44	si:dkey-243k1.3	myl9b	cldnb	ptprc	si:ch211-270g19.5	gamt	manf	si:dkey-126g1.9	nupr1	ftr82	icn	plpp1a
45	si:ch211-255i20.3	sepp1a	si:dkey-16l2.20	nfkbie	aplnra	leg1.2	calm1b	grapa	si:dkey-243k1.3	krt8	myct1a	plvap
46	pabpc4	MMP23B	zgc:162730	ptges3a	tgfb3	ahsg1	atp6v0cb	mmel1	CR753876.1	actb2	rpl6	cavin1b
47	si:dkey-266f7.9	angptl1a	hsp90ab1	chordc1a	pfn2	c9	pgm3	vamp5	rtn1a	pdzk1ip1	pros1	krt18
48	gamt	rpl36	pdlim1	srgn	si:ch211-1a19.3	rnasel2	bhlhe40	aqp1a.1	gamt	si:ch73-31d8.2	rps28	actb2
49	nupr1	rps2	slc25a5	lcp1	csrp1b	afp4.1	nkx3.2	tmsb4x	si:dkey-266f7.9	si:ch73-86n18.1	gas6	ptprb
50	CR753876.1	rpl37.1	kcnk5b	ptpn6	gadd45ba	rbp2b	npm1a	higd1a	cel.1	c5	rpl11	anxa13

	Pancreas endocrine cells (gip)	Peripheral Neurons	Macrophages	Hepatic duct cells	Apelin expressing cells	Pancreas duct epithelial cells	Pancreas endocrine cells (epsilon)	Bi-hormonal endocrine cells	Granulocytes	Pancreas duct epithelial cells	T cells	Hepatic duct cells
# cells	264	259	255	168	155	151	97	84	44	33	32	30
# of av. UMIs	18146	3006	4475	3761	1217	7996	5864	1218	3013	1449	2249	3187
# of av. genes	1398	695	970	989	389	1318	878	434	733	496	575	1053
top 50 genes	gip	scn4ab	grn1	si:ch1073-406l10	hbegfb	anxa4	mlnl	TMEM269	lect2l	tnfrsf11b	si:ch211-212d10.1	rdh10a
2	si:ch211-56a11.2	clic6	cd74a	cart3	fgl2a	lgals2b	ghrl	zbtb16a	lyz	tm4sf4	il4	COLEC10
3	atp1b1b	scn1ba	si:busm1-266f07.2	si:ch73-256j6.5	apln	epcam	mboat4	CT027638.1	BX908782.3	anxa4	il13	steap4
4	si:dkey-1h6.8	sncga	c1qb	slco2b1	pim2	tm4sf4	pdyn	zgc:158463	cebp1	prr15la	nfkbiaa	colec11
5	pcp4b	calm1a	ctsd	vcam1b	gp1bb	si:ch211-137i24.10	BX908782.1	cacna2d2a	ncf1	hopx	laptm5	angptl6
6	hpgd	cldn11a	c1qc	ccbe1	itga2b	si:ch211-153b23.5	plcxd3	qyllt1b	npsn	cldnb	si:dkey-21e2.13	plp1b
7	kctd12.1	entpd3	cd74b	si:ch211-106h4.12	hspa4a	si:dkey-96g2.1	clu	mast2	mmp13a	zgc:158258	MFAP4 (1 of many).8	lrata
8	BX908782.1	tmem176	tspan36	CU570881.2	hsp70.3	spint2	glula	si:ch1073-266p11.2	scpp8	zgc:172079.2	MFAP4 (1 of many).1	hhip
9	foxa2	marcksl1b	fcer1gl	tlx1	srgn	cldnc	cst3	gcga	cxcr4b	gpx1b	tnfrsf9a	ecm1b
10	tango2	gpm6ab	lygl1	rgs5b	hsp70l	elf3	si:rp71-17i16.6	kcnma1a	pfn1	si:dkey-61f9.2	si:dkey-21e2.15	bmp2a
11	hepacam2	fabp7a	si:dkey-5n18.1	pcsk6	CR383676.1	fam213b	trdn	CR847844.1	il6r	bik	MFAP4 (1 of many).9	igfbp5b
12	gadd45aa	itgb4	c1qa	cxcl18b	il11b	cfl1l	abhd17c	arl4ab	nccrp1	si:ch73-208h1.1	hspa8	agtr2
13	admb	cd59	psap	agtr2	pfn1	si:ch73-359m17.9	zgc:92818	ins	mpx	si:ch73-22a13.3	ccr9a	si:dkey-42i9.16
14	ulk1a	nrgna	ccl35.1	krt18	admb	gpx1b	acsl1a	cacna1aa	fcer1gl	grtp1a	si:dkey-21e2.10	ccbe1
15	capn2b	nr4a2b	marco	grem2b	rap1b	cdh17	CR847984.1	gcgb	si:dkey-10p5.7	si:dkeyp-73b11.8	tnfrsf9b	rspo3
16	pdx1	calm3a	atp6v0ca	tnw	hsp70.1	cd9b	tmsb1	ddx61	CR318653.2	tnks1bp1	traf1	tbp5
17	zgc:101731	prnpb	ms4a17a.10	si:ch73-256j6.4	si:ch211-105j21.9	cldn15a	brd1b	prodha	si:ch73-248e21.5	tmem176l.4	pfn1	ackr4b
18	zgc:65894	si:dkey-23i12.7	zgc:92066	nr1h5	pdcd4b	cldn7b	hepacam2	selo	srgn	lgals2b	ccl34b.4	bgnb
19	sh3gl2a	hspb1	mfap4	c6.1	blf	si:rp71-17l1.1	tm4sf4	sst1.1	si:ch211-264f5.2	KRT18 (1 of many).1	stip1	ntn1a
20	scg3	wfdc2	grn2	tfpia	hacd3	zgc:64022	nat8l	parp2	hspa8	si:ch211-79k12.1	srgn	phex
21	calm1b	rgcc	scpep1	s100u	hsp70.2	zgc:198329	scg5	hhla2a.2	lta4h	elf3	hspa4a	pcolcea
22	rgs5a	calm2a	ctss2.1	TPM1	rsrp1	krt18	ca6	hspa9	si:ch73-343l4.8	sdc4	MFAP4 (1 of many).7	aoc2
23	mid1ip1b	si:ch211-137a8.4	ctss2.2	arhgdig	si:ch211-214p13.9	phlda2	snap25a	ero1b	rac2	mcl1a	DNAJA4	calb2a
24	scg2a	cbx7a	ctsz	igfbp3	chac1	slc5a1	anxa4	CR383676.1	arpc1b	adgrg11	chchd2	glula
25	syngr3b	id1	lgmn	msx1b	hsp90aa1.2	cdh1	stmn1b	zgc:194659	si:ch1073-429i10.1	ric8a	si:ch211-112c15.8	tmem88b
26	copz2	cd81a	pfn1	chodl	si:ch211-222i21.1	hbegfa	nupr1	foxo1a	CR848841.2	cldnc	hsp90aa1.2	nr1h5
27	mllt11	junba	si:ch211-212k18.7	myl9b	dusp5	ccl19b	pttg1ipb	BX571707.1	alox5ap	krt5	si:dkey-21e2.14	rpz3
28	syt4	phlda2	sqstm1	taqln	qcnt4a	slc6a14	adh8a	abcc5	hspa4a	arq2	hspbp1	rpz5
29	aldocb	hspb8	CD53	rdh10a	si:ch211-202a12.4	ftr82	asah1b	fam46c	bik	hsp90b1	hspa5	si:ch1073-291c23.2
30	vqf	icn	spi1b	tp63	si:ch211-195b11.3	tspan35	zgc:172120	kmt2ca	s1pr4	wu:fb59d01	si:dkey-112a7.4	sfrp1a
31	npm1a	efhd1	laptm5	serpinf1	fermt3b	hbegfb	viml	tnrc6c1	zgc:77112	ankrd22	cremb	si:ch1073-406l10.2
32	neurod1	htra1a	cx32.2	tmem176l.1	DNAJA4	slc25a5	hrh2b	jph1a	stip1	agr2	VSIR	si:ch73-29l19.1
33	rrad	serpinh1b	cxcr4b	sdc4	tln1	si:ch211-213d14.2	ehbp1l1a	zbtb16b	si:ch211-284o19.8	si:dkey-61f9.1	hspe1	cd82a
34	insm1b	hsp70.3	ccl35.2	tmsb4x	thbs1b	cx30.3	nucb2b	si:dkey-177p2.6	si:dkey-19b23.14	calm1b	wasb	rgs4
35	nupr1	hsp90aa1.2	NPC2 (1 of many)	pfn1	atp2b1a	rassf7b	insm1a	fam13b	capgb	adgrf3a	si:dkey-21e2.16	marcksl1b
36	egr4	hsp70l	lgals2a	AEBP1 (1 of many)	grb2b	zgc:193726	plp1b	chtopa	lpar5a	hsp70.2	hsp90b1	zgc:110340
37	chchd10	ip6k2a	cd63	fosb	grasp	icn	scgn	hsp70.1	coro1a	si:dkey-33i11.9	CR383676.1	sepp1a
38	snap25a	sparc	cycsb	actb2	plek	si:ch211-139a5.9	ndrg2	gfra1a	pnp5a	klf6a	il11b	masp1
39	fev	uchl1	havcr1	zgc:153952	tal1	pdzk1ip1	pcsk1nl	camk2n1a	si:ch73-248e21.7	il17rc	erh	c1qtnf5
40	mtbl	glula	spi1a	jdp2b	Sep 12	serpinb1	neurod1	PPP2CB (1 of many)	gpr183a	hepacam2	aldocb	bmp6
41	ppdpfb	jdp2b	zgc:109934	krt8	bcl2l10	cnn2	camk2n1a	si:ch73-376l24.2	sepw1	wbp2	ets2	si:dkey-238o13.4
42	ak4	gulp1a	si:cabz01074946.1	eps8l1	atf4b	adh8b	fybb	sik2b	cfl1l	mknk2b	bzw1b	si:zfos-464b6.2
43	gadd45ga	cish	arpc1b	fhl2a	zfand2a	cldnb	mllt11	hsph1	itb4r	herpud1	rgs13	tmem176
44	hmp19	pnrc2	si:dkey-27i16.2	zgc:195173	bc2	mibp2	cplx2l	CABZ01020261.1	si:dkey-102g19.3	hspa5	bcl10	tmem176l.1
45	zgc:153845	higd1a	anxa3b	acta2	CD63	actb2	rcan2	slc20a1b	plekhf1	f11r.1	CR847844.1	si:dkey-8k3.2
46	rasd1	tspan35	vmp1	CU207311.2	ubb	zgc:162730	ccni	cct7	ch25hl2	krt18	phactr2	tsc22d3
47	tuba1c	hsp70.2	rac2	clic2	hspa8	chia.6	hmp19	tbcb	cotl1	epcam	si:ch211-214j24.10	cyp3c1
48	NA	fosb	coro1a	myh11a	arpc1b	cldnh	npas4a	bhlhe40	tmsb4x	rhogb	arpc1b	sparc
49	NA	mcl1b	MFAP4 (1 of many).4	myl6	tmsb4x	idh1	egr2b	eif5	si:ch1073-443f11.2	stip1	RASA2	zgc:195173
50	NA	zfp36l1b	grna	tpm4a	dnajb1b	si:ch73-335l21.4	egr4	mid1ip1b	DNAJA4	sqstm1	gata3	zgc:113372

Additional table 4.

Cell clusters and differentially expressed genes for the whole zebrafish heart.

Cell clusters and differentially expressed genes for the whole zebrafish heart at 0, 3, 7 and 30 dpi

diff. genes (top 50)	B-cells	Endothelial cells (apnln)	Endothelial cells (lyve1)	Endothelial cells (plvapb)	Cardiomyocytes (dediff.)	Cardiomyocytes (Atrium)	Cardiomyocytes (Ventricle)	Endocardium (Atrium)	Endocardium (frzb)	Endocardium (Ventricle)	Epicardium (Atrium)
1	ighv1-4	hbegfb	cxcl12a	wu:fj16a03	nppa	tnnc1b	tnni4a	zgc:158343	vwf	si:ch73-86n18.1	gstm.3
2	CU571382.1	admb	CU929150.1	cxcl12b	nppb	myh6	CR926459.1	spock3	si:ch211-153b23.5	aqp8a.1	s100a10a
3	igic1s1	apln	thy1	rbp2a	hsp90aa1.1	si:ch211-270g19.5	fabp3	im:7152348	c2cd4a	spock3	fn1b
4	si:dkey-234i14.9	si:ch211-195b11.3	lyve1a	plvapb	ttn.2	tcap	myl7	ptgs2a	edn2	mb	krt15
5	CU896602.1	sele	selenop	cldn5b	ttn.1	tnni1b	ak1	ptgs2b	cfd	fabp11a	mmp2
6	zgc:153659	itga2b	cdh6	rgcc	myh6	aldoab	cmlc1	id2b	frzb	epas1b	si:ch211-105c13.3
7	igl3v5	thbs1b	si:dkey-203a12.9	fabp11a	atp2a2a	tnnt2a	actc1a	vcam1b	ecscr	ramp2	krt4
8	zgc:194275	fgl2a	lyve1b	id1	xirp1	smtnl1	actа1b	aqp8a.1	glulb	si:ch211-145b13.6	frzb
9	CR318588.1	FO681357.1	id1	tmsb1	si:ch211-131k2.3	cmlc1	myh7l	mycb	efemp2b	f8	mmel1
10	mgst3b	pim2	si:ch211-105c13.3	ldb2a	hspb11	gapdh	ndufa4	her6	calm1a	her6	anxa2a
11	ccr9a	gp1bb	arl4aa	tcima	her4.1	tnnc1a	gapdh	krt18	cyr61	cdh5	ppdpfa
12	BX571825.5	cldn5b	si:ch211-152c2.3	cavin1b	mybpc3	nppa	ckma	grb10a	nr4a2b	fosab	krt94
13	cxcr4b	gcnt4a	hapln3	rgs5b	lmo7a	ddx54	eno3	clic2	cpamd8	si:ch211-248e11.2	rbp4
14	ccl35.1	ilf1b	stab1	hopx	si:ch211-270g19.5	ckma	tnnt2a	si:ch211-248e11.2	nr4a1	eppk1	si:ch211-286o17.1
15	CR936442.1	si:ch211-105j21.9	loxl1	cdh5	unc45b	zgc:101840	atp5pd	gpr182	cdh5	socs3a	anxa1a
16	cxcr4a	tln1	ramp2	kdrl	tcap	mybphb	cox5b2	klf2a	s100v2	tmem88b	CR318588.3
17	si:busm1-266f07.2	si:ch211-214p13.9	etv2	scarb2a	ryr2b	myl7	cox7a1	cxcl18b	appa	podxl	dkk3b
18	grap2b	f5	hsd3b7	gig2j	tnnc1b	tpm4a	aldoaa	fosl1a	cyr61l2	cyp1b1	tcf21
19	rgs13	si:ch211-214p16.2	ctsla	bcl6b	synpo2lb	actc1a	tnnc1a	tmem88b	hbegfa	grb10a	postnb
20	cd74b	si:ch211-153b23.5	rgcc	bsg	ldb3a	acta1b	tnni1b	dusp5	si:ch211-248e11.2	si:ch73-335i21.4	slc29a1b
21	CT573342.2	si:ch211-214p16.1	akap12a	podxl	pla1a	idh2	tpm4a	glulb	entpd8	lpl	tmsb1
22	CT573231.1	rap1b	tppp3	mcamb	sorbs1	cox5b2	cox4i1l	si:ch73-335i21.4	her9	jun	phactr4a
23	si:dkey-24p1.1	blf	tcima	pdgfbb	bves	mpc1	mdh2	akap12b	klf2a	si:ch211-160j14.3	podxl
24	cd74a	sep12	ch25h	jam2a	flnca	mdh2	ppdpfa	ctsla	ptmaa	spon1b	c6.1
25	b2m	myh9a	fabp11a	ecscr	ACTC1	uqcrfs1	mustn1b	fosb	ptgs2a	tmem88a	myh10
26	dapp1	fam212ab	CR762483.1	si:dkey-126g1.9	slc8a1a	mdh1aa	atp5mc1	ackr4b	krt15	akap12b	tuba1a
27	zgc:171713	plek	plvapb	krt18	ndrg4	idh2	igf2b	ehd2b	klf2b	klf2b	f3b
28	tspan2a	lpp	sox7	ptprb	trdn	ak1	nme2b.2	bzw1b	crip2	id2b	timp2a
29	BX005450.1	srgn	qstt1b	epas1b	myh7l	ptp4a3	rbpms2b	si:ch211-153b23.5	spock3	zgc:64106	tbx18
30	si:ch1073-220m6.1	pfn1	pde4ba	si:ch211-156j16.1	pdk2a	nmrk2	pgam2	tmem88a	lmo2	egfl7	endouc
31	ponzr6	hsp70l	lgmn	si:dkey-262k9.2	pabpc4	atp5mc3b	ldb3b	zfp36l1b	junba	ecscr	cxcl14
32	BX571825.3	hacd3	plk2b	CABZ01030107.1	myom1b	atp5mc1	hspb11	ramp2	s100a10b	tspan4b	ckba
33	si:ch211-150o23.2	si:ch211-250g4.3	hyal2b	plvapa	cox7c	calm3a	coxbb1	errfi1a	id2b	fam174b	aldh1a2
34	si:busm1-194e12.11	fermt3b	mrc1b	sox7	hspb8	atp5pd	atp5meb	ponzr1	clec14a	kdrl	col6a1
35	ncf1	blvrb	rab11bb	cav1	cited4b	atp5meb	cox6a2	fosab	icn	junba	c1qtnf5
36	hspbp1	atp2b1a	gpm6ab	ch25h	ATP5MD	atp5f1b	ATP5MD	si:ch211-145b13.6	clic2	ackr4b	hsd11b2
37	hspa4a	cbx7a	stab2	plpp3	desma	chchd10	mpc1	btg2	cav1	zgc:152791	c7a
38	rplp2l	hsp70.1	dab2	col4a1	cox6a2	adprhl1	gamt	junba	gpd1b	si:dkey-261h17.1	si:ch211-156j16.1
39	si:ch211-119e14.1	hsp70.2	myct1a	CABZ01075125.1	zgc:101840	fhl2a	atp5po	cyp2ad2	fosl2	si:dkeyp-97a10.2	id2a
40	cd83	crema	zfp36l1b	gpm6ab	mt-co1	ckbb	atp5mc3b	nppc	tie1	junbb	rn7sk
41	blnk	zgc:171686	ponzr1	fgd5a	bag3	atp5mc3a	zgc:193541	lpl	qadd45ba	nppc	acvrl1
42	DNAJA4	mpl	flt4	pecam1	ITGB1BP2	lmod2b	tpi1b	slc22a31	errfi1a	slc22a31	spaca4l
43	rps17	tpte	clec14a	ca16b	kcnh6a	si:dkey-51e6.1	uqcr10	sat1a.2	BX663503.3	klf6a	mxra8b
44	dusp2	lmo2	ebf3a.1	gpm6aa	tnni1b	cox5aa	slc25a5	jun	sox7	clic2	mfap5
45	si:ch211-195b13.1	arpc1b	s1pr1	cavin2b	tnnt2a	ndufb10	csrp3	junbb	net1	appa	s100a10b
46	atp11a	grasp	carhsp1	TCIM (1 of many)	laptm4b	mybpc3	atp5if1b	tmem2	ackr4b	flt1	c4
47	bcl6a	hsp70.3	fxyd6l	limch1b	smtnl1	zgc:85722	atp5f1d	vwf	serpine1	tfpia	lxn
48	coro1a	myh11a	cldn11a	igfbp1a	aldoab	slc25a3b	atp5f1e	kctd12.2	sult2st3	ier2b	flnb
49	pdcd4b	si:ch211-222l21.1	ms4a17a.8	clec14a	pygmb	atp5if1b	cox7b	spint2	CABZ01030107.1	im:7152348	olfml3b
50	syk	bmp16	igf2a	oaz2a	cox7a1	hbaa1	mdh1aa	pim1	efnb2a	ccdc187	cav1

diff. genes (top 50)	Epicardium (Ventricle)	Valve Fibroblasts	Fibroblasts	Macrophages	Monocytes	Myelin cells	Neuronal cells	Neutrophils	Perivascular cells	Proliferating cells	Smooth muscle cells	T-cells
1	si:ch211-106h4.12	zgc:153704	col1a2	grn1	ccl35.1	scn4ab	vipb	lyz	TCIM (1 of many)	pcna	rgs5a	si:ch211-214p16.1
2	gstm.3	abi3bpb	mdka	grn2	epdl1	cd59	elavl4	lect2l	cxcl12b	stmn1a	acta2	ccl34b.4
3	krt15	krt4	col1a1a	ccl35.1	si:ch211-214p16.1	mbpa	stmn1b	BX908782.2	pdgfrb	hmgb2b	TPM1 (1 of many)	ccl36.1
4	postnb	angptl7	col1a1b	cd74a	si:busm1-266f07.2	mpz	syt1a	npsn	rasl12	DUT	C11orf96	ccl38.6
5	s100a10a	igfbp5b	col5a1	c1qb	si:ch211-214p16.2	mbpb	tuba1c	si:ch211-117m20.5	BX901920.1	sumo3b	tagln	DNAJA4
6	fn1b	cyp26b1	dpt	ctsd	ccl35.2	plp1b	gap43	mmp13a.1	kcne4	rrm2.1	myh11a	ccr9a
7	endouc	tnfaip6	fn1b	c1qc	CU459094.3	BX936284.1	snap25a	si:ch211-9d9.1	rgs5b	si:ch211-288g17.3	krt91	CR936442.1
8	mmp2	mgp	ccl25b	lygl1	cd74a	anxa13l	stmn2a	scpp8	rgs4	rpa3	itih1	cxcr4b
9	tfa	rspo1	sparc	lgals3bpb	arpc1b	apoeb	sncga	si:ch1073-67j19.1	aqtr2	banf1	myl6	hspbp1
10	krt4	fibinb	clu	si:ch211-147m6.2	cd74b	si:rp71-19m20.1	rtn1b	mmp9	TPM1 (1 of many)	rpa2	myl9a	zgc:64051
11	krt94	aif1l	dcn	marco	tmsb1	sncga	mllt11	pnp5a	si:ch211-198c19.3	lig1	BX663503.3	rgs13
12	frzb	krt91	mfap5	fcer1gl	anxa3b	nrgna	crhbp	mpx	ifitm1	fen1	vim	pfn1
13	si:ch211-105c13.3	bhmt	postnb	ctss2.2	sftpbb	stmn1b	slc5a7a	ncf1	fxyd6l	sfrp5	si:dkey-164f24.2	nkl.1
14	cldnc	gyg1b	c4	mfap4	id2a	itgb4	anxa13l	fcer1gl	slc20a1a	dnajc9	ctgfa	pdcd4b
15	tmsb1	si:dkey-4p15.3	fn1a	NPC2 (1 of many)	mpeg1.1	scn1ba	zgc:85894	pfn1	calm2b	rbbp4	lmod1b	hspa4a
16	si:ch211-198c19.3	serpinh1b	col12a1a	si:ch211-147m6.1	mhc2dab	fstl3	stmn2b	scinlb	notch3	tyms	flna	sla2
17	adh8a	si:dkey-205h13.1	c6.1	cd74b	cxcr4b	cldnk	ddc	cfl1l	cd248a	rrm1	cfd	cebpb
18	si:ch1073-459j12.1	tuft1a	si:ch1073-459j12.1	grna	si:dkey-33i11.9	sox10	chata	cotl1	tagln	top2a	tpm2	cyb5a
19	mdka	f13a1b	htra1b	si:busm1-266f07.2	zgc:92066	CABZ01068367.1	gng3	plekhf1	si:ch73-193j22.1	nutf2l	si:busm1-57f23.1	coro1a
20	c6.1	ITGB1BP2	pmp22a	ctss2.1	pfn1	pdgfbb	FO907089.1	cpa5	col18a1b	cks1b	zgc:77517	stip1
21	anxa2a	crip1	hapln1a	atp6v0ca	hspbp1	si:ch211-286c4.6	phox2a	illr4	mcamb	hells	mdkb	b2m
22	col1a2	carhsp1	ckba	hmox1a	DNAJA4	nanos1	pcsk1nl	myh9a	atp1a1b	chaf1a	tns1b	si:dkey-27i16.2
23	podxl	uchl1	col6a3	zgc:174904	si:dkey-27i16.2	entpd3	scg2b	adam8a	stk17al	selenoh	si:ch211-1a19.3	si:dkey-109a10.2
24	si:ch211-137i24.10	fbln5	gstm.3	si:dkey-5n18.1	stip1	timd4	atp1b2a	si:dkey-102g19.3	rgs5a	mki67	mylka	si:ch211-202a12.4
25	col1a1a	krt15	col6a1	ctsba	cxcl11.6	lim2.2	cpe	lta4h	mustn1a	zgc:110540	gch2	chchd2
26	spaca4l	rbp5	tnfaip6	rasgef1ba	aif1l	foxd3	nsg2	timp2b	myh11a	mcm5	slmapb	hsp90aa1.2
27	igfbp5b	gch2	tfa	lgmn	hspa4a	acbd7	vip	il6r	pcdh18a	mcm6	cyr61	rhbdd1
28	col1a1b	pcolcea	mmp2	c1qa	nr4a3	pou3f1	vgf	si:ch73-54b5.2	plp1b	ubr7	lpp	cxcr4a
29	rbp4	irs2b	f3b	psap	synqr3b	cldn19	hpcal4	cfbl	fhl3b	si:ch211-156b7.4	dnajb4	tnfb
30	eppk1	cbx7a	rbp4	sqstm1	sqstm1	uqt8	atp1a3a	vsir	sh3d21	smc2	CR855337.1	calm1b
31	sparc	snorc	col5a2a	CR855311.1	zgc:152863	gldn	cplx2	alox5ap	ccdc3b	si:ch211-66k16.27	fhl1a	tnfrsf9b
32	anxa1a	hspb8	CR936442.1	ctsz	rhbdd1	si:ch211-197g15.7	phox2bb	nccrp1	prx	mcm4	fhl3b	ptprc
33	dkk3b	rpl22l1	ifitm1	pfn1	hspe1	fabp2	prph	arg2	snai1a	mcm3	fbln5	si:ch211-105j21.9
34	cygb1	thbs1b	sostdc1a	atp6v1g1	BX004816.2	znf536	syngr	si:ch211-284o19.8	sparc	rrm2	alcama	dusp2
35	col6a3	ptx3a	anxa2a	si:ch211-212k18.7	pdcd4b	ninj2	map1b	CU499336.2	si:ch211-251b21.1	cdk1	tmsb4x	arpc1b
36	tgm2b	timp4.3	pcolcea	tspan36	zgc:103700	depdc7	maptb	si:ch1073-443f1.2	si:ch211-270g19.5	dek	c2cd4a	sh2d1ab
37	si:dkey-8k3.2	csrp2	olfml3b	ctsc	chchd2	si:ch211-137a8.4	rab6bb	s1pr4	slc7a2	rpa1	htra1a	FP236356.1
38	sostdc1a	laptm4b	tmsb2	si:dkey-27i16.2	actb2	tlcd1	tuba2	arpc1b	myo1b	h2afx	phlda2	wasb
39	mgp	cebpd	si:ch211-105c13.3	mrc1b	zdhhc18b	elovl1a	syt11a	DNAJA4	jam3a	mcm2	anxa1a	BX120005.1
40	aldh1a2	crip2	si:ch211-137i24.10	cd9b	psap	si:ch211-234p6.5	chrna3	anxr2a	arhgap36	nasp	cav2	capgb
41	c4	mmp2	c1qtnf5	zgc:92066	rasgef1ba	col15a1b	stxbp1a	si:ch211-136m16.8	AL954322.2	cks2	cst3	si:dkey-262k9.2
42	crip1	dcn	ppib	ccl34a.4	ctss2.2	egr2b	si:ch73-119p20.1	tmsb1	BX908750.1	gins2	zgc:92162	si:ch211-165d12.4
43	zgc:123068	sec61g	si:ch211-106h4.12	tppp	zgc:64051	gpm6ab	ret	si:ch73-343l4.8	atp1b4	asf1ba	f3a	BX005105.1
44	ppdpfa	rps27.2	col6a2	cotl1	spi1b	fstb	vamp2	srgn	en1b	dhfr	sort1a	laptm5
45	hsd11b2	tgm2l	mgp	ncf1	zgc:136870	nr4a2b	cplx2l	zgc:77112	CABZ01068367.1	prim2	cald1b	oser1
46	fn1a	GSTZ1	cxcl12a	cfp	coro1a	dhrs3b	cntnap2a	coro1a	oscp1a.1	rfc5	kng1	mcl1a
47	col5a1	matn4	edil3a	laptm5	hspa8	phlda2	sypa	stip1	ppp1r14aa	rnasehh2a	epas1a	wbp2
48	col6a1	hmgb3a	dkk3b	coro1a	hsp90aa1.2	calm1a	eef1a1a	cd7al	rftn2	chtf8	wfdc2	fkbp4
49	colec12	MYO1D	si:ch211-198c19.3	si:ch211-194m7.4	wbp2	cd9b	stx1b	si:ch73-248e21.5	rassf4	si:dkey-6i22.5	tpm4b	scinlb
50	si:ch211-286o17.1	fxyd6l	fstl1b	scpep1	ckbb	vwa1	elavl3	CT030188.1	spdya	dctd	fhl2a	si:dkey-177p2.6

Additional table 5.

Sub-clustering and differentially expressed genes for *col1a1a* expressing cells during regeneration.

diff. genes (top 50)	Epicardium (Atrium)	Epicardium (Ventricle)	Fibroblasts (const.)	Fibroblasts (cfd)	Fibroblasts (col11a1a)	Fibroblasts (col12a1a)	Fibroblasts (cxcl12a)	Fibroblasts (mpeg1.1)	Fibroblasts (nppc)	Fibroblasts (proliferating)
1	s100a10a	si:ch211-106h4.12	hapln1a	cfd	tnc	col12a1a	adh8a	grn1	nppc	ube2c
2	CR318588.3	cldnc	myl7	ccl25b	fn1a	serpine1	si:dkey-8k3.2	ccl35.1	si:ch211-248e11.2	kpna2
3	mmel1	eppk1	cmlc1	si:ch1073-406l10.2	col11a1a	mdka	clu	cd74a	aqp8a.1	top2a
4	ppdpfa	tmsb1	fgl2b	sfrp1b	col12a1a	serpinh1b	ccl25b	grn2	nppb	cdk1
5	gstm.3	podxl	si:ch211-5k11.8	rgs5a	col5a1	mfap2	slco2b1	fcer1gl	serpine1	mki67
6	CR318588.1	krt15	hbba1.1	ccl20b	col1a2	postnb	cxcl12a	c1qb	clic2	pcna
7	slc29a1b	endouc	NC-002333.4	pcolcea	col1a1b	lgals2a	dhrs3a	c1qc	sele	cks1b
8	podxl	gstm.3	dpt	clu	timp2b	timp2b	rbp4	cd74b	vwf	cenpf
9	tmsb1	krt94	mt-nd3	plpp1a	col11a1b	col5a1	c4	lgals3bpb	spock3	col11a1a
10	si:ch211-286o17.1	s100a10a	hbaa1	C11orf96	col1a1a	sparc	selenop	si:ch211-147m6.2	cdh5	stmn1a
11	phactr4a	cavin1b	hbba1	mfap5	tnfaip6	si:busm1-57f23.1	tfa	ctss2.2	zgc:158343	DUT
12	myh10	krt91	mfap5	ifitm1	fkbp9	col11a1b	hsd11b2	si:busm1-266f07.2	lepb	mad2l1
13	frzb	adh8a	htra1b	cxcl12a	lrrc17	lum	pcolcea	si:ch211-147m6.1	si:ch211-153b23.5	aspm
14	krt94	aldh1a2	CR936442.1	si:dkey-164f24.2	col5a2a	nupr1	bhmt	ctss2.1	ptgs2a	rrm2.1
15	CR318588.4	frzb	hbba2	cyp2ad2	vmp1	col1a2	vmo1a	marco	rhoca	zgc:194627
16	cav1	krt8	gapdh	cyp1c2	adam8a	CR753876.1	lxn	grna	zfand2a	ccnb1
17	zgc:158404	tmem88b	tnnt2a	lmod1b	col12a1b	col1a1a	dpt	mfap4	ctsla	si:ch211-69g19.2
18	anxa1a	spaca4l	CR753876.1	mylka	pcolce2b	col12a1b	tmem176	zgc:64051	DNAJA4	cdc20
19	itm2cb	postnb	si:dkey-261h17.1	dpt	zgc:85975	vwa1	rspo3	NPC2 (1 of many)	vmp1	nusap1
20	tuba1a	cav1	sostdc1a	acta2	asph	DNAJA4	CR855311.1	c1qa	krt18	si:ch211-266i6.3
21	aldh1a2	ppdpfa	si:ch211-137i24.10	ccl19a.1	cthrc1a	ppib	c6.1	zgc:174904	kcne4	rpa2
22	krt15	sost	NC-002333.17	lum	ckba	fn1a	cemip	coro1a	crlf1a	cks2
23	si:ch211-66e2.3	cfl1l	pmp22a	rbp4	tagln	col1a1b	si:ch1073-291c23.2	cxcr4b	tfpia	anln
24	id2a	anxa1a	actc1a	mdkb	mdka	CR936442.1	steap4	si:dkey-5n18.1	fabp11a	plk1
25	calm1a	si:ch211-198c19.3	hbaa2	osr1	fkbp7	hspa5	ponzr1	laptm5	hspb8	tp53inp2
26	cxcl14	c3a.1	tnnc1a	ponzr1	cygb1	sulf2b	pmp22a	ncf1	edn2	tpx2
27	zgc:152791	cygb1	phlda2	si:ch1073-459j12.1	efemp2a	edil3a	cpn1.1	rac2	hspa5	marcksb
28	cavin2b	si:ch211-105c13.3	atf5b	ctgfa	cxcl12a	aebp1	mgst1.2	ptprc	sat1a.2	lmnb2
29	arvcfb	slc43a3b	nppa	mgst1.2	serpinh1a	rgcc	stmn1b	cfp	ramp2	si:dkey-30c15.10
30	tmem98	tmem98	tnni1b	rasl11b	postnb	si:dkey-12l12.1	ntn1a	mrc1b	dusp5	kif23
31	si:ch211-105c13.3	gpx1a	tnnc1b	vmo1a	fkbp11	clu	c6	csf3b	errfi1a	cdca8
32	ecm1b	cemip	ckma	tnfsf12	lgals2a	nid1b	si:ch211-106h4.12	si:dkey-102g19.3	id2b	rpa3
33	krt4	si:ch211-286o17.1	fam107b	tagln	rcn3	tnfaip6	cxcl18b	mpeg1.1	hspb1	birc5a
34	mao	zgc:110182	myh7l	c6.1	kdelr2b	col5a2a	stmn1a	tppp	hsp90aa1.2	si:ch211-156b7.4
35	gata5	CR925728.1	acta1b	dcn	ssr3	pcolce2b	aldh9a1a.1	si:ch211-194m7.4	pnp5a	kif20a
36	tcf21	bnc2		tsc22d3	tuba8l3	zgc:92161	tsc22d3	zgc:153317	tspan36	ccna2
37	sox6	mmp2		thbs4b	calua	fstl1b	prss23	mafbb	glulb	prdx1
38	atoh8	cyp2ad3		glula	loxa	col6a1	cebpd	ccl34a.4	egfl7	smc2
39	si:ch211-250c4.4	eps8l3b		olfml3b	ddost	snai2	serpinf1	mhc2dab	arl4ab	aurkb
40	cavin1b	si:ch211-137i24.10		si:busm1-57f23.1	eif5a2	col6a2	soul5	zgc:173915	hspbp1	spc24
41	smad6a	efna1b		TPM1 (1 of many)	ssr2	tmsb2	mfap2	pglyrp5	sult2st3	kif11
42	gcga	phactr4a		myh11a	tubb5	hsp90b1	thbs3a	si:ch211-102c2.4	si:ch211-160j14.3	dek
43	si:dkey-188i13.7	si:ch211-250c4.4		si:ch211-132p1.3	txn	mxra8a	fabp11a	spi1b	ahsa1b	cenpx
44	wnt11r	krt4		arl5c	eno1a	ctsk	gchfr	slc2a6	sox7	tyms
45	lrrc15	igfbp5b		b2m	rrbp1b	si:dkey-27i16.2	mdkb	ms4a17a.10	atp1b1a	CABZ01058261.1
46	tbx18	cotl1		cd74b	cox7a2a	zgc:85975	rgs5b	zgc:103700	hspa4a	prc1b
47	si:ch211-156j16.1	CR318588.3		saa	hint1	plxdc2	ifitm1	slc7a7	prelid3b	tacc3
48	frmd8	CR318588.4		zgc:195173	gapdhs	col8a1b	ptgdsb.2	stoml3b	sik1	chaf1a
49	igfbp2a	aldh9a1a.1		tpm2	dad1	dpt		tcirg1b	srgn	fen1

diff. genes (top 50)	Fibroblasts (spock3)	Valve Fibroblasts	Perivascular cells
1	spock3	zgc:153704	TCIM (1 of many)
2	aqp8a.1	abi3bpb	cxcl12b
3	epas1b	angptl7	TPM1 (1 of many)
4	si:ch211-145b13.6	rspo1	BX901920.1
5	cdh5	cyp26b1	rgs5a
6	si:ch211-248e11.2	aif1l	rasl12
7	grb10a	si:dkey-205h13.1	si:ch211-270g19.5
8	vwf	igfbp5b	myh11a
9	zgc:64106	si:ch211-131k2.3	rgs4
10	ecscr	f13a1b	agtr2
11	fam174b	gyg1b	pdgfrb
12	egfl7	gch2	kcne4
13	gpr182	bhmt	fabp11a
14	clec14a	ITGB1BP2	marcksl1b
15	f8	krt4	slc20a1a
16	yrk	si:dkey-4p15.3	oaz2a
17	lpl	cbx7a	cd248a
18	fli1a	carhsp1	notch3
19	id2b	fibinb	mcamb
20	sox7	fbln5	stk17al
21	tmem88a	irs2b	si:ch73-193i22.1
22	mb	snorc	atp1a1b
23	errfi1a	tuft1a	plp1b
24	fabp11a	nr4a3	pcdh18a
25	edn2	itih1	mylka
26	dll4	nr4a2b	sh3d21
27	marcksl1b	hsp90aa1.2	ccdc3b
28	slc22a31	DNAJA4	slc7a2
29	si:ch211-153b23.5	rps27.2	BX908750.1
30	ackr4b	eef2l2	jam3a
31	si:ch73-86n18.1	mcl1a	arhgap36
32	nrarpa	csrp2	cdkn1ca
33	nr4a1	matn4	atp1b4
34	ramp2	CR855337.1	en1b
35	ponzr1	sec61g	ppp1r14aa
36	fosab	vcanb	oscp1a.1
37	im:7152348	tgm2l	rassf4
38	appa	pnrc2	sorbs2b
39	dusp5	hsp70l	spdya
40	hsp70l	si:ch211-200p22.4	pkib
41	fosl1a	si:dkey-19b23.8	ndufa4l2a
42	her6	cyr61l2	ncaldb
43	clic2	slc38a4	tagln
44	vcam1b	MYO1D	prx
45	fosb	rpl22l1	acta2
46	mycb	wwp2	rftn2
47	glulb	rsl24d1	calm2b
48	klf2a	si:ch211-152c2.3	rasgef1ba
49	si:ch73-335l21.4	timp4.3	rbpms2a
50	zgc:158343	rps26	atp2a3

Additional table 6.

Probes used for in-situ hybridization.

Gene Name	Accession Number	Company	Catalog #
col1a1a	NM_199214.1	ACD	409491
col11a1a	XM_005162814.1	ACD	803311-C3
col12a1a	XM_002665259.6	ACD	556481-C2
nppc	NM_001109940.1	ACD	556501-C3
cxcl12a	NM_178307.2	ACD	406481-C2
s100a10a	NM_001005961.2	ACD	556491-C2
itm2cb	NM_199980.1	ACD	556521-C3
ttn.2	XM_021479070	ACD	810421
mpeg1.1	NM_212737.1	ACD	536171-C3
pdgfrb	NM_001190933.1	ACD	493921-C2
angptl7	NM_001006073	ACD	845191-C3
cyp26b1	NM_212666.1	ACD	571281-C2
aldh1a2	NM_131850.1	ACD	455681-C3
pcna	NM_131404.2	ACD	574931-C3
mCherry	N/A	ACD	513201

Additional table 7.

Primer sequences.

	Primers for sgRNA synthesis
sgRNA primer 1	TAATACGACTCACTATAGGTGTCCACGTAGTAGTAGCGTTTTAGAGCTA GAAATAGCAAG
sgRNA constant oligo	AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCT TATTTTAACTTGCTATTTCTAGCTCTAAAAC
	Primers for bulk scar detection
	(forward primers differ only in their barcode sequence)
Reverse primer	GTTCAGAGTTCTACAGTCCGACGATCGTCCTCGTTGTGGGAGGTG
Forward primer 1	CCTTGGCACCCGAGAATTCCACTGGTCGGCGGCACGCTGATCTACAAGG
Forward primer 2	CCTTGGCACCCGAGAATTCCAAGGCTCTACGGCACGCTGATCTACAAGG
Forward primer 3	CCTTGGCACCCGAGAATTCCAGTGGTGCCCGGCACGCTGATCTACAAGG
Forward primer 4	CCTTGGCACCCGAGAATTCCAACGATGTCCGGCACGCTGATCTACAAGG
Forward primer 5	CCTTGGCACCCGAGAATTCCACCGATTGACGGCACGCTGATCTACAAGG
Forward primer 6	CCTTGGCACCCGAGAATTCCAAATTAACCCGGCACGCTGATCTACAAGG
Forward primer 7	CCTTGGCACCCGAGAATTCCACATTACAACGGCACGCTGATCTACAAGG
Forward primer 8	CCTTGGCACCCGAGAATTCCAATTGACCGCGGCACGCTGATCTACAAGG
Forward primer 9	CCTTGGCACCCGAGAATTCCATGTCACGACGGCACGCTGATCTACAAGG
Forward primer 10	CCTTGGCACCCGAGAATTCCAGCTAACTGCGGCACGCTGATCTACAAGG
Forward primer 11	CCTTGGCACCCGAGAATTCCACTTGAGACCGGCACGCTGATCTACAAGG
Forward primer 12	CCTTGGCACCCGAGAATTCCATCTAAGGCCGGCACGCTGATCTACAAGG
Forward primer 13	CCTTGGCACCCGAGAATTCCACGTCATAGCGGCACGCTGATCTACAAGG
Forward primer 14	CCTTGGCACCCGAGAATTCCAACTAATCACGGCACGCTGATCTACAAGG
Forward primer 15	CCTTGGCACCCGAGAATTCCATATTATGGCGGCACGCTGATCTACAAGG
Forward primer 16	CCTTGGCACCCGAGAATTCCATATGCACGCGGCACGCTGATCTACAAGG
	Primers for single cell scar detection (see also the protocol from 10X Genomics Chromium)
RFP specific Fwd primer 1	CGGCACGCTGATCTACAAGG
RFP specific Fwd primer 2	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGAGTTCAAGACCATC TACATGGCC
Reverse primer	CTACACGACGCTCTTCCGATCT

Acknowledgement

I would like to thank

Dr. Bastiaan Spanjaard

Sara Lelek

for the close collaboration on the projects, and

Nina Mitic

Jana Richter

Ronny Schäfer

for their support with the experiments, and

Karoline Holler

Anika Neuschulz

Gaurav Jumde

Nora Fresmann

Dr. Roberto Moreno-Ayala

Dr. Pedro Olivares

for the discussions, and

Dr. Daniela Panakova

Dr. Holger Gerhardt

Dr. Nikolaus Rajewsky

for conceptual supervision, and special thank to

Dr. Jan Philipp Junker

For providing me the opportunity and daily supervision to perform this thesis.